

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:32:03 ; Search time 12.5 Seconds  
(without alignments)  
184.688 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAGAAERPGEAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	100.0	309	2	A39773
2	115	100.0	332	2	A38873
3	111	96.5	309	2	A39169
4	111	96.5	335	2	S08341
5	104	90.4	281	2	A41400
6	50	43.5	399	2	AC3089
7	50	43.5	399	2	G98197
8	48	41.7	261	2	AH0089
9	48	41.7	1058	2	T50496
10	47	40.9	352	2	T36719
11	46.5	40.4	1293	2	T01512
12	46	40.0	250	2	T45028
13	46	40.0	317	2	A87479
14	46	40.0	416	1	W2MLDP
15	45	39.1	405	2	T47595
16	45	39.1	409	2	S44346
17	45	39.1	421	2	E87277
18	45	39.1	923	2	A39596
19	45	39.1	1088	2	C83054
20	44.5	38.7	281	2	JC4558
21	44.5	38.7	335	2	F84411
22	44.5	38.7	811	2	T36581
23	44	38.3	194	2	A70609
24	44	38.3	287	2	A54601
25	44	38.3	346	2	A55111
26	44	38.3	346	2	JC7638
27	44	38.3	432	2	D86937
28	44	38.3	486	2	C75154
29	44	38.3	503	2	T35053

30 44 38.3 515 2 C75615 expolyposphatase  
31 44 38.3 1207 2 T23754 hypothetical prote  
32 43.5 37.8 661 2 C81822 topoisomerase IV s  
33 43.5 37.8 661 2 D81055 60s acidic ribosom  
34 43 37.4 110 1 R6BYP3 hypothetical prote  
35 43 37.4 118 2 S27476 histone H1 homolog  
36 43 37.4 154 2 T04159 LIM protein - rat  
37 43 37.4 327 2 JC4385 Ig alpha-2 chain C  
38 43 37.4 340 1 A2HU mannose-6-phosphat  
39 43 37.4 391 2 JQ1192 mannose-6-phosphat  
40 43 37.4 391 2 AC0691 sugar-binding prot  
41 43 37.4 401 2 AE3583 involucrin - ring-  
42 43 37.4 450 1 A43733 glycine hydroxymet  
43 43 37.4 574 2 G71336 mucin TIM-C.1 - Af  
44 43 37.4 562 2 A45155 hypothetical prote  
45 43 37.4 584 2 T40319

#### ALIGNMENTS

##### RESULT 1

A39773  
myristylated alanine-rich protein kinase C substrate, macrophage - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 21-Feb-1992 #sequence revision 21-Feb-1992 #text\_change 18-Feb-2000  
C:Accession: A39773; S29267; S05681  
R:Brusallinsky, J.D.; Brooks, S.F.; Herget, T.; Morris, C.; Rozengurt, E.  
J. Biol. Chem. 266, 7073-7080, 1991  
A:Title: Molecular cloning and characterization of the acidic 80-kDa protein kinase C sub  
A:Reference number: A39773; MUID:91201362; PMID:1707878  
A:Accession: A39773  
A>Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-309 <ERU1>  
A:Cross-references: GB:M59859; NID:g205068  
A:Note: This translation is not annotated in GenBank entry RATKINC, release 113.0  
R:Herget, T.; Brooks, S.F.; Broad, S.; Rozengurt, E.  
Eur. J. Biochem. 209, 7-14, 1992  
A:Title: Relationship between the major protein kinase C substrates acidic 80-kDa protein  
or equivalent genes in different species.  
A:Reference number: S29267; MUID:93011168; PMID:1396720  
A:Accession: S29267  
A:Molecule type: mRNA  
A:Residues: 182,'E',184-301 <HER>  
R:Brusallinsky, J.D.; Morris, C.; Perks, K.; Brown, R.; Brooks, S.; Rozengurt, E.  
FEBS Lett. 255, 149-153, 1989  
A:Title: Internal amino acid sequence analysis of the 80 kDa protein kinase C substrate  
A:Reference number: S05681; MUID:90005952; PMID:2676596  
A:Accession: S05681  
A:Molecule type: protein  
A:Residues: 12-28, GX, 31,189-215,'E',217,252-275,'X',299-309 <ERU2>  
A:Note: the amino end of the mature protein is blocked  
C:Superfamily: neurofilament triplet H protein  
C:Keywords: blocked amino end; lipoprotein; myristylation; phosphoprotein  
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted

Query Match 100.0%; Score 115; DB 2; Length 309;  
Best Local Similarity 100.0%; Pred. No. 4.1e-09;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGAAERPGEAAVA 24

Db 2 GAQFSKTAAGAAERPGEAAVA 25

##### RESULT 2

A38873  
myristylated alanine-rich protein kinase C substrate - human  
N:Alternate names: acidic calmodulin-binding 80K protein; MARCKS  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text\_change 20-Jun-2000  
C:Accession: A38873; A42977; A40758; S29269

R;Shimizu, N.  
 submitted to DDBJ, September 1991  
 A;Reference number: A38873  
 A;Accession: A38873  
 A;Molecule type: mRNA  
 A;Residues: 1-332 <SHI>  
 A;Cross-references: GB:D10522; GB:D90498; NID:g219893; PIDN:BAA01392.1; PID:g219894  
 R;Sakai, K.; Hirai, M.; Kudoh, J.; Minoshima, S.; Shimizu, N.  
 Genomics 14, 175-178, 1992  
 A;Title: Molecular cloning and chromosomal mapping of a cDNA encoding human 80K-L protein  
 A;Reference number: A42977; MUID:93052291; PMID:1427823  
 A;Accession: A42977  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-13, 'G', 15-332 <SAK>  
 A;Cross-references: GB:D90498  
 A;Experimental source: squamous carcinoma cells A431  
 A;Note: sequence extracted from NCBI backbone (NCBI:P:118653)  
 R;Harlan, D.M.; Graff, J.M.; Stumpo, D.J.; Eddy Jr., R.L.; Shows, T.B.; Boyle, J.M.; Bla  
 J. Biol. Chem. 266, 14399-14405, 1991  
 A;Title: The human myristoylated alanine-rich C kinase substrate (MARCKS) gene (MACS). A  
 A;Reference number: A40758; MUID:913117795; PMID:1860846  
 A;Accession: A40758  
 A;Molecule type: mRNA  
 A;Residues: 1-83, 'A', 85-118, 'P', 120-233, 'W', 235-286, 'LVC', 290, 'RRGSGRGGARRSLNQ', 309-3  
 A;Cross-references: GB:M68956  
 A;Note: the authors translated the codon GGC for residue 53 as Arg  
 R;Herget, T.; Brooks, S.F.; Broad, S.; Rozengurt, E.  
 Eur. J. Biochem. 209, 7-14, 1992  
 A;Title: Relationship between the major protein kinase C substrates acidic 80-kDa protei  
 or equivalent genes in different species.  
 A;Reference number: S29267; MUID:93011168; PMID:1396720  
 A;Accession: S29269  
 A;Molecule type: mRNA  
 A;Residues: 189-223, 'R', 225-234, 'E', 236-322 <HER>  
 C;Comment: This protein is a major cellular substrate for protein kinase C and plays a r  
 C;Genetics:  
 A;Gene: GDB:MACS  
 A;Cross-references: GDB:118835; OMIM:177061  
 A;Map position: 6q22.2-6q22.2  
 C;Superfamily: neurofilament triplet H protein  
 C;Keywords: actin binding; blocked amino end; calmodulin binding; lipoprotein; myristyla  
 F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 P;159,163,167,170/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status

Query Match 100.0%; Score 115; DB 2; Length 332;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-09;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAARPGEAAVA 24  
 |||||  
 DB 2 GAQFSKTAAGGAAARPGEAAVA 25

RESULT 3  
 A39169  
 myristylated alanine-rich protein kinase C substrate, macrophage - mouse  
 N;Alternate names: MARCKS; myristoylated alanine-rich C kinase substrate  
 C;Species: Mus musculus (house mouse)  
 C;Date: 24-Jan-1992 #sequence\_revision 24-Jan-1992 #text\_change 18-Feb-2000  
 A;Accession: A39169; S16519; S29268; S63977; S74153  
 R;Seykora, J.T.; Ravetch, J.V.; Aderem, A.  
 Proc. Natl. Acad. Sci. U.S.A. 88, 2505-2509, 1991  
 A;Title: Cloning and molecular characterization of the murine macrophage "68-kDa" protei  
 A;Reference number: A39169; MUID:91172836; PMID:2006186  
 A;Accession: A39169  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-309 <SEY>  
 A;Cross-references: GB:M60474; NID:g199026; PIDN:AAA39491.1; PID:g199027  
 R;Brooks, S.F.; Herget, T.; Erusalimsky, J.D.; Rozengurt, E.  
 EMBO J. 10, 2497-2505, 1991

A;Title: Protein kinase C activation potentially down-regulates the expression of its major  
 A;Reference number: S16519; MUID:91330872; PMID:1868832  
 A;Accession: S16519  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-95, 'T', 97, 'T', 99-309 <EMB>  
 R;Herget, T.; Brooks, S.F.; Broad, S.; Rozengurt, E.  
 Eur. J. Biochem. 209, 7-14, 1992  
 A;Title: Relationship between the major protein kinase C substrates acidic 80-kDa protei  
 or equivalent genes in different species.  
 A;Reference number: S29267; MUID:93011168; PMID:1396720  
 A;Accession: S29268  
 A;Molecule type: mRNA  
 A;Residues: 182-301 <HER>  
 R;Herget, T.; Oehrlein, S.A.; Pappin, D.J.C.; Rozengurt, E.; Parker, P.J.  
 Eur. J. Biochem. 233, 448-457, 1995  
 A;Title: The myristoylated alanine-rich C-kinase substrate (MARCKS) is sequentially phosph  
 A;Reference number: S63977; MUID:96067684; PMID:7586787  
 A;Accession: S63977  
 A;Molecule type: protein  
 A;Residues: 145-151, 'X', 153-155, 'X', 157-162, 'X', 164-169 <HEW>  
 R;Schoenwasser, D.C.; Palmer, R.H.; Herget, T.; Parker, P.J.  
 FEBS Lett. 395, 1-5, 1996  
 A;Title: p42 MAPK phosphorylates 80 kDa MARCKS at Ser-113.  
 A;Reference number: S74153; MUID:97002278; PMID:8849678  
 A;Accession: S74153  
 A;Molecule type: protein  
 A;Residues: 102-130 <SCH>  
 C;Superfamily: neurofilament triplet H protein  
 C;Keywords: actin binding; blocked amino end; calmodulin binding; lipoprotein; myristyl  
 F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 P;152,156,160,163/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status

Query Match 96.5%; Score 111; DB 2; Length 309;  
 Best Local Similarity 95.8%; Pred. No. 1.5e-08;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAARPGEAAVA 24  
 |||||  
 DB 2 GAQFSKTAAGGAAARPGEAAVA 25

RESULT 4  
 S08341  
 myristylated alanine-rich protein kinase C substrate - bovine  
 N;Alternate names: acidic calmodulin-binding 81K protein (ACAMP-81); MARCKS  
 C;Species: Bos primigenius taurus (cattle)  
 C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 18-Feb-2000  
 A;Accession: S08341; A32904; S29270; A46098; PS0338  
 R;Stumpo, D.J.; Graff, J.M.; Albert, K.A.; Greengard, P.; Blackshear, P.J.  
 Nucleic Acids Res. 17, 3987-3988, 1989  
 A;Title: Nucleotide sequence of a cDNA for the bovine myristoylated alanine-rich C kinase  
 A;Reference number: S08341; MUID:89282412; PMID:2734111  
 A;Accession: S08341  
 A;Status: translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-335 <STU>  
 A;Cross-references: EMBL:M24638; NID:g163339; PIDN:AAA30635.1; PID:g163340  
 R;Stumpo, D.J.; Graff, J.M.; Albert, K.A.; Greengard, P.; Blackshear, P.J.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 4012-4016, 1989  
 A;Title: Molecular cloning, characterization, and expression of a cDNA encoding the "80-  
 A;Reference number: A32904; MUID:89264553; PMID:2726763  
 A;Accession: A32904  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-98, 'Q', 100-335 <ST2>  
 A;Cross-references: GB:M24638; GB:M23738  
 R;Herget, T.; Brooks, S.F.; Broad, S.; Rozengurt, E.  
 Eur. J. Biochem. 209, 7-14, 1992  
 A;Title: Relationship between the major protein kinase C substrates acidic 80-kDa protei  
 or equivalent genes in different species.  
 A;Reference number: S29267; MUID:93011168; PMID:1396720  
 A;Accession: S29270

A:Molecule type: DNA  
 A:Residues: 191-253, 'SEE', 257-279, 283-292, 'V', 294, 'PQOE', 299, 'A', 300, 'A', 302-313, 'A', 315  
 R:Manenti, S.; Sorokine, O.; Van Dorsselaer, A.; Taniguchi, H.  
 J. Biol. Chem. 268, 6878-6881, 1993  
 A:Title: Isolation of the non-myristoylated form of a major substrate of protein kinase  
 A:Reference number: A46098; MUID:93216617; PMID:8463217  
 A:Accession: A46098  
 A:Molecule type: protein  
 A:Residues: 2-11 <VAN>  
 R:Mizutani, A.; Tokumitsu, H.; Hidaka, H.  
 Biochem. Biophys. Res. Commun. 182, 1395-1401, 1992  
 A:Title: Acidic calmodulin binding protein, ACAMP-81, is MARCKS protein interacting with  
 A:Reference number: PS0338; MUID:92171958; PMID:1540183  
 A:Accession: PS0338  
 A:Molecule type: protein  
 A:Residues: 12-30; 56-69; 88-98, 'AS', 100-103; 104-109, 'E', 111-123; 156-160; 165-171; 196-215; 2  
 A:Experimental source: brain  
 A:Note: This sequence is identical with that of myristoylated alanine-rich C-kinase sub  
 C:Comment: This protein is a major cellular substrate for protein kinase C and plays a r  
 C:Comment: It binds to calmodulin in one to one molar ratio in the presence of calcium a  
 C:Genetics:  
 A:Introns: 34/3  
 C:Superfamily: neurofilament triplet H protein  
 C:Keywords: actin binding; blocked amino end; calmodulin binding; lipoprotein; myristyla  
 F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 F:158,162,166,169/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status  
 Query Match 96.5%; Score 111; DB 2; Length 335;  
 Best Local Similarity 95.8%; Pred. No. 1.6e-08;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GQAQFSKTAAGGAAARPGGAAVA 24  
 |||||  
 DB 2 GQAQFSKTAAGGAAARPGGAAVA 25  
 |||||  
 RESULT 5  
 A41400  
 myristylated alanine-rich protein kinase C substrate - chicken  
 C:Species: Gallus gallus (Chicken)  
 C:Date: 03-Apr-1992 #sequence\_revision 03-Apr-1992 #text\_change 21-Jul-2000  
 C:Accession: A41400  
 R:Graff, J.M.; Stumpo, D.J.; Blackshear, P.J.  
 Mol. Endocrinol. 3, 1903-1906, 1989  
 A:Title: Molecular cloning, sequence, and expression of a cDNA encoding the chicken myri  
 A:Reference number: A41400; MUID:90114197; PMID:2608063  
 A:Accession: A41400  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-281 <GRA>  
 A:Cross-references: GB:M31650; NID:g212287; PIDN:AAA48946.1; PID:g212288  
 C:Superfamily: neurofilament triplet H protein  
 C:Keywords: actin binding; blocked amino end; lipoprotein; myristylation; phosphoprotein  
 F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted  
 Query Match 90.4%; Score 104; DB 2; Length 281;  
 Best Local Similarity 87.5%; Pred. No. 1.4e-07;  
 Matches 21; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 GQAQFSKTAAGGAAARPGGAAVA 24  
 |||||  
 DB 2 GQAQFSKTAAGGAAARPGGAAVA 25  
 |||||  
 RESULT 6  
 AC3089  
 conserved hypothetical protein Atu4335 [imported] - Agrobacterium tumefaciens (strain C5  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
 C:Accession: AC3089  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I  
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
 i Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I  
 ser, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; MUID:21608550; PMID:11743193  
 A:Accession: AC3089  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-399 <KUR>  
 A:Cross-references: GB:AE008689; PIDN:AAL45129.1; PID:g17742801; GSPDB:GN00187  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atu4335  
 A:Map position: linear chromosome  
 Query Match 43.5%; Score 50; DB 2; Length 399;  
 Best Local Similarity 50.0%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
 QY 3 QFSKTAAGGAAARPGGAAVA 24  
 |||||  
 DB 367 QFLQRAARGAAASARETDATAA 388  
 |||||  
 RESULT 7  
 G98197  
 hypothetical protein AGR\_L\_1057 [imported] - Agrobacterium tumefaciens (strain C58, Cerec  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 18-Nov-2002  
 C:Accession: G98197  
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;  
 Science 294, 2323-2328, 2001  
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
 A:Reference number: A97359; MUID:21608551; PMID:11743194  
 A:Accession: G98197  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-399 <KUR>  
 A:Cross-references: GB:AE007870; PIDN:AAK89105.1; PID:g15158909; GSPDB:GN00170  
 C:Genetics:  
 A:Gene: AGR\_L\_1057  
 A:Map position: linear chromosome  
 Query Match 43.5%; Score 50; DB 2; Length 399;  
 Best Local Similarity 50.0%; Pred. No. 12;  
 Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
 QY 3 QFSKTAAGGAAARPGGAAVA 24  
 |||||  
 DB 367 QFLQRAARGAAASARETDATAA 388  
 |||||  
 RESULT 8  
 AH0089  
 probable flagellar basal-body rod protein flgG [imported] - Yersinia pestis (strain CO92)  
 C:Species: Yersinia pestis  
 C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Nov-2001  
 C:Accession: AH0089  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
 geno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.; I  
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, I  
 Nature 413, 523-527, 2001  
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
 A:Reference number: AB0001; MUID:21470413; PMID:11586360  
 A:Accession: AH0089  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-261 <KUR>  
 A:Cross-references: GB:AL590842; PIDN:CAC89579.1; PID:g15978811; GSPDB:GN00175  
 C:Genetics:  
 A:Gene: flgG  
 C:Superfamily: rod protein flgF





Search completed: October 7, 2004, 17:36:01  
Job time : 16.5 secs

A:Molecule type: DNA  
A:Residues: 1-317 <STO>  
A:Cross-references: GB:AE005673; NID:gl3423295; PIDN:AAK23829.1; GSFDB:GNO0148  
C:Genetics:  
A:Gene: CCL854

Query Match 40.0%; Score 46; DB 2; Length 317;  
Best Local Similarity 50.0%; Pred. No. 37;  
Matches 12; Conservative 2; Mismatches 10; Indels 0; Gaps 0;  
QY 1 GAQFSKTAAGGAAERPGGAAYA 24  
Db 27 GARISWCAAKGETVYTGELKAVVA 50

RESULT 14  
W2WLDP  
E2 protein - deer papillomavirus  
C:Species: deer papillomavirus  
A:Note: host Odocoileus virginianus (American white-tailed deer)  
C:Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 24-Feb-1994  
C:Accession: A03673  
R:Groff, D.E.; Lancaster, W.D.  
J Virol. 56; 85-91, 1985  
A:Title: Molecular cloning and nucleotide sequence of deer papillomavirus.  
A:Reference number: A93013; MUID:85293253; PMID:2993669  
A:Accession: A03673  
A:Molecule type: DNA  
A:Residues: 1-416 <GRO>  
C:Superfamily: papillomavirus E2 protein  
C:Keywords: early protein

Query Match 40.0%; Score 46; DB 1; Length 416;  
Best Local Similarity 60.0%; Pred. No. 47;  
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 9 AKGEAAERPGGAAY 23  
Db 368 AVGEQGSERPGDATV 382  
RESULT 15  
T47595  
RING finger protein T12E18.50 - Arabidopsis thaliana  
N:Alternate names: protein T12E18.50  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 05-May-2000  
C:Accession: T47595  
R:Blocker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat M.Mewes,  
submitted to the Protein Sequence Database, March 2000  
A:Reference number: Z24469  
A:Accession: T47595  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-405 <BLO>  
A:Cross-references: EMBL:AL132971  
A:Experimental source: cultivar Columbia; BAC clone T12E18  
C:Genetics:  
A:Gene: T12E18.50  
A:Map position: 3  
A:Introns: 130/2; 208/3; 257/2; 286/3; 335/3; 385/3  
C:Superfamily: Arabidopsis thaliana RING finger protein T12E18.50; RING finger homology  
F104-148/Domain: RING finger homology <RRN>

Query Match 39.1%; Score 45; DB 2; Length 405;  
Best Local Similarity 37.5%; Pred. No. 64;  
Matches 9; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAERPGGAAYA 24  
Db 30 GCPFSKAARPPDDASARKQETTAS 53

**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:32:03 ; Search time 8 seconds  
(without alignments)  
156,210 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GQFSKTAAGGAARPGAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query %	Score	Match	Length	ID	Description
1	115	100.0	308	1	MACS_RAT	P30009 rattus norv
2	115	100.0	331	1	MACS_HUMAN	P29966 homo sapien
3	111	96.5	308	1	MACS_MOUSE	P26645 mus musculu
4	111	96.5	331	1	MACS_BOVIN	P12624 bos taurus
5	104	90.4	280	1	MACS_CHICK	P16527 gallus gall
6	67	58.3	85	1	MRXD_RAT	P20468 rattus norv
7	46	40.0	382	1	OXAA_PROMA	Q7V800 prochloroco
8	46	40.0	416	1	VE2_FAPVD	P03123 deer papill
9	45	39.1	409	1	R23B_HUMAN	P54727 homo sapien
10	45	39.1	923	1	PRGR_MOUSE	Q00175 mus musculu
11	45	39.1	1587	1	TOP2_PENCH	Q9V868 penicillium
12	44	38.3	291	1	USP2_RAT	Q63655 rattus norv
13	44	38.3	346	1	USP2_HUMAN	Q15853 homo sapien
14	44	38.3	346	1	USF2_MOUSE	Q64705 mus musculu
15	43	37.4	110	1	RLA3_SCHPO	P17477 schizosacch
16	43	37.4	118	1	YNIF_AZOB	P25316 azospirillu
17	43	37.4	246	1	BRM1_MOUSE	Q99n20 mus musculu
18	43	37.4	306	1	TEP1_MOUSE	O54819 mus musculu
19	43	37.4	326	1	PD11_MOUSE	O70400 mus musculu
20	43	37.4	327	1	PD11_RAT	P25294 rattus norv
21	43	37.4	391	1	MANA_SALTY	P25081 salmonella
22	43	37.4	421	1	SP39_BRUAB	O06875 brucella ab
23	43	37.4	421	1	SP39_BRUME	Q8YCE2 brucella me
24	43	37.4	421	1	SP39_BRUSU	Q8FXV5 brucella su
25	43	37.4	450	1	INVO_LEMCA	P14590 lemur catta
26	43	37.4	463	1	FMNL_HUMAN	O95466 homo sapien
27	43	37.4	574	1	GLYA_TREPA	O83349 treponema p
28	43	37.4	662	1	MUC1_XENLA	Q05049 xenopus lae
29	43	37.4	684	1	ISH1_SCHPO	Q9V7X6 schizosacch
30	43	37.4	1013	1	A60D_DROME	P91327 drosophila
31	43	37.4	1322	1	PUR4_XYLFA	Q9PDI6 xylella fas
32	43	37.4	2339	1	CCAB_RABIT	Q05152 cryptotagus
33	42	36.5	261	1	RL1_HSV2H	P28283 herpes simp

34	42	36.5	262	1	FLGG_CAUCR	Q06172 caulobacter
35	42	36.5	274	1	BPAL_STRAU	P33912 streptomyce
36	42	36.5	351	1	CSP_PLAKU	P04922 plasmodium
37	42	36.5	442	1	CYS4_DICDI	P54639 dictyosteli
38	42	36.5	495	1	AB31_CHLRE	Q8S339 chlamydomon
39	42	36.5	627	1	MUTL_RHILO	Q98312 rhizobium l
40	42	36.5	645	1	DNAK_RHOBA	Q7um31 rhodopirell
41	42	36.5	733	1	HIC1_HUMAN	Q14526 homo sapien
42	42	36.5	788	1	NASP_HUMAN	P49321 homo sapien
43	42	36.5	1159	1	SOR2_HUMAN	Q96pq0 homo sapien
44	42	36.5	1198	1	HCN4_RAT	Q9Jka7 rattus norv
45	41	35.7	106	1	RLA3_YEAST	P10622 saccharomyc

#### ALIGNMENTS

##### RESULT 1

ID	MACS_RAT	STANDARD;	PRT;	308 AA.
AC	P30009;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Myristoylated alanine-rich C-kinase substrate (MARCKS).			
GN	MARCKS OR MACS.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=91201362; PubMed=1707878;			
RA	Brusalmisky J.D., Brooks S.F., Herzog T., Morris C., Rozengurt E.;			
RT	"Molecular cloning and characterization of the acidic 80-kDa protein			
RT	kinase C substrate from rat brain. Identification as a			
RT	glycoprotein.";			
RL	J. Biol. Chem. 266:7073-7080(1991).			
RN	[2]			
RP	PHOSPHORYLATION SITES.			
RC	TISSUE=Brain;			
RX	MEDLINE=93135774; PubMed=8422248;			
RA	Heemskerk F.M., Chen H.C., Huang F.L.;			
RT	"Protein kinase C phosphorylates Ser-152, Ser-156 and Ser-163 but not			
RT	Ser-160 of MARCKS in rat brain.";			
RL	Biochem. Biophys. Res. Commun. 190:236-241(1993).			
CC	-!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR			
CC	PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND			
CC	SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.			
CC	-!- PTM: PHOSPHORYLATION BY PKC REPLACES MARCKS FROM THE MEMBRANE. IT			
CC	ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.			
CC	-!- SIMILARITY: Belongs to the MARCKS family.			

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
EMBL; M59859; -; NOT\_ANNOTATED\_CDS.

PIR; A39773; A39773.

InterPro; IPR002101; MARCKS.

Pfam; PF02063; MARCKS; 1.

PRINTS; PR00963; MARCKS.

PROSITE; PS00826; MARCKS 1; 1.

PROSITE; PS00827; MARCKS 2; 1.

Phosphorylation; Myristate; Calmodulin-binding; Actin-binding;

Membrane; Lipoprotein.

INIT MET 0 0 BY SIMILARITY.

LIPID 1 1 N-myristoyl glycine (By similarity).

```
FT DOMAIN 144 168 CALMODULIN-BINDING (PSD) .
FT MOD_RRS 151 151 PHOSPHORYLATION (BY PKC) .
FT MOD_RRS 155 155 PHOSPHORYLATION (BY PKC) .
FT MOD_RRS 162 162 PHOSPHORYLATION (BY PKC) .
SQ SEQUENCE 308 AA; 29663 MW; 59E50228BB1B75D2 CRC64;

Query Match 100.0%; Score 115; DB 1; Length 308;
Best Local Similarity 100.0%; Pred. No. 2.3e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GAQFSKTAAGGAAAAERPGAAVA 24
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GAQFSKTAAGGAAAAERPGAAVA 24

RESULT 2
MACS_HUMAN
ID MACS_HUMAN STANDARD; PRT; 331 AA.
AC P29966;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Myristoylated alanine-rich C-kinase substrate (MARCKS) (Protein kinase
DE C substrate, 80 kDa protein, light chain) (PKCSL) (80K-L protein).
GN MARCKS OR MACS OR PRKCSL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91317795; PubMed=1860846;
RA Haylan D.M., Graff J.M., Stumpo D.J., Eddy R.L. Jr., Shows T.B.,
RA Boyle J.M., Blackshear P.J.;
RT "The human myristoylated alanine-rich C kinase substrate (MARCKS)
RT gene (MACS). Analysis of its gene product, promoter, and chromosomal
RT localization."
RL J. Biol. Chem. 266:14399-14405(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93052291; PubMed=1427823;
RA Sakai K., Hirai M., Kudoh J., Minoshima S., Shimizu N.;
RT "Molecular cloning and chromosomal mapping of a cDNA encoding human
RT 80K-L protein: major substrate for protein kinase C."
RL Genomics 14:175-178(1992).
CC -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR
CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND
CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.
CC -!- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT
CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.
CC -!- SIMILARITY: Belongs to the MARCKS family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M68956; AAA59555.1; -
CC DR EMBL; M68955; AAA59554.1; -
CC DR EMBL; D10522; BAA01392.1; -
CC DR PIR; A38873; A38873.
CC DR Genew; HGNC:6759; MARCKS.
CC DR MIM; 177061; -
CC DR GO; GO:0035629; C:actin cytoskeleton; TAS.
CC DR GO; GO:0003780; F:actin cross-linking activity; TAS.
CC DR GO; GO:0005516; F:calmodulin binding; TAS.
CC DR InterPro; IPR002101; MARCKS.
CC DR Pfam; PF02063; MARCKS; 1.
CC DR PRINTS; PR00963; MARCKS.
CC DR PROSITE; PS00826; MARCKS_1; 1.
```

```
DR PROSITE; PS00827; MARCKS_2; 1.
KW Phosphorylation; Myristate; Calmodulin-binding; Actin-binding;
KW Membrane; Lipoprotein.
FT INIT_MET 0 0 BY SIMILARITY.
FT LIPID 1 1 N-myristoyl glycine (By similarity).
FT DOMAIN 151 175 CALMODULIN-BINDING (PSD).
FT MOD_RRS 158 158 PHOSPHORYLATION (BY PKC).
FT MOD_RRS 162 162 PHOSPHORYLATION (BY PKC).
FT MOD_RRS 166 166 PHOSPHORYLATION (BY PKC).
FT MOD_RRS 169 169 PHOSPHORYLATION (BY PKC).
FT CONFLICT 83 83 S -> A (IN REF. 1).
FT CONFLICT 118 118 A -> P (IN REF. 1).
FT CONFLICT 233 233 P -> S (IN REF. 1).
FT CONFLICT 286 307 PGAPPEQEAFAAEFAAAASS -> LVCPRRGGSPRGAR
GT GRSLLQ (IN REF. 1).
SQ SEQUENCE 331 AA; 31413 MW; BCC837D586581774 CRC64;

Query Match 100.0%; Score 115; DB 1; Length 331;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GAQFSKTAAGGAAAAERPGAAVA 24
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GAQFSKTAAGGAAAAERPGAAVA 24

RESULT 3
MACS_MOUSE
ID MACS_MOUSE STANDARD; PRT; 308 AA.
AC P26645;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myristoylated alanine-rich C-kinase substrate (MARCKS).
GN MARCKS OR MACS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Macrophage;
RX MEDLINE=91172836; PubMed=2006186;
RA Seykora J.T., Ravetch J.V., Aderem A.;
RT "Cloning and molecular characterization of the murine macrophage '68-
RT kDa' protein kinase C substrate and its regulation by bacterial
RT lipopolysaccharide."
RL Proc. Natl. Acad. Sci. U.S.A. 88:2505-2509(1991).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX TISSUE=Fibroblast;
RX MEDLINE=91330872; PubMed=1868832;
RA Brooks S.F., Herget T., Erusalimsky J.D., Rozengurt E.;
RT "Protein kinase C activation potentially down-regulates the expression
RT of its major substrate, 80K, in Swiss 3T3 cells."
RL EMBO J. 10:2497-2505(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein N.J., Udels T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McWeney P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
```

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmitt J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RL human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [4]  
 RP PARTIAL SEQUENCE.  
 RC STRAIN=Swiss; TISSUE=Fibroblast;  
 RX MEDLINE=90346162; PubMed=2384168;  
 RA Brooks S.F., Brusalsky J.D., Totty N.F., Rozengurt E.,  
 RT "Purification and internal amino acid sequence of the 80 kDa protein  
 RT kinase C substrate from Swiss 3T3 fibroblasts. Homology with  
 RT substrates from brain.";  
 RL FEBS Lett. 268:291-295 (1990).  
 CC -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR  
 CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND  
 CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.  
 CC -!- TISSUE SPECIFICITY: BRAIN, SPLEEN, LESS IN KIDNEY AND HEART, AND  
 CC VERY LOW LEVELS IN LIVER.  
 CC -!- INDUCTION: BY lipopolysaccharides (LPS).  
 CC -!- PTM: PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT  
 CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.  
 CC -!- SIMILARITY: Belongs to the MARCKS family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M60474; AAA39491.1; -;  
 DR EMBL; BC046601; AAA46601.1; -;  
 DR PIR; A39169; A39169.  
 DR MGD; MGI:96907; Marcks.  
 DR InterPro; IPR002101; MARCKS.  
 DR Pfam; PF02063; MARCKS.  
 DR PRINTS; PR00963; MARCKS.  
 DR PROSITE; PR00826; MARCKS\_1; 1.  
 DR PROSITE; PS00827; MARCKS\_2; 1.  
 KW Phosphorylation; Myristate; Calmodulin-binding; Actin-binding;  
 KW Membrane; Lipoprotein.  
 FT INIT MET 0  
 FT LIPID 1  
 FT DOWAIN 144 168  
 FT MOD RES 151 151  
 FT MOD RES 155 155  
 FT MOD RES 159 159  
 FT MOD RES 162 162  
 FT MOD RES 162 162  
 FT CONFLICT 95 97  
 FT AGA -> TGT (IN REF. 2).  
 SQ SEQUENCE 308 AA; 29530 MW; FB5313B913701CSC CRC64;  
 Query Match 96.5%; Score 111; DB 1; Length 308;  
 Best Local Similarity 95.8%; Pred. No. 8.8e-09;  
 Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 GAQFSKTAAGKGAARERGEAAVA 24  
 Db 1 GAQFSKTAAGKGAARERGEAAVA 24  
 RESULT 4  
 ID MACS\_BOVIN STANDARD; PRT; 331 AA.  
 AC P12624;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Myristoylated alanine-rich C-kinase substrate (MARCKS) (ACAMP-81).  
 GN MARCKS OR MACS.

OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89282412; PubMed=2734111;  
 RA Stumpo D.J., Graff J.M., Albert K.A., Greengard P., Blackshear P.J.,  
 RT "Nucleotide sequence of a cDNA for the bovine myristoylated  
 RT alanine-rich C kinase substrate (MARCKS).";  
 RL Nucleic Acids Res. 17:3987-3988 (1989).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=89264553; PubMed=2726763;  
 RA Stumpo D.J., Graff J.M., Albert K.A., Greengard P., Blackshear P.J.,  
 RT "Molecular cloning, characterization, and expression of a cDNA  
 RT encoding the '80- to 87-kDa' myristoylated alanine-rich C kinase  
 RT substrate: a major cellular substrate for protein kinase C.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:4012-4016 (1989).  
 RN [3]  
 RP PARTIAL SEQUENCE.  
 RX MEDLINE=92171958; PubMed=1540183;  
 RA Mizutani A., Tokumitsu H., Hidaka H.,  
 RT "Acidic calmodulin binding protein, ACAMP-81, is MARCKS protein  
 RT interacting with synapsin I.";  
 RL Biochem. Biophys. Res. Commun. 182:1395-1401 (1992).  
 RN [4]  
 RP PHOSPHORYLATION SITES.  
 RX MEDLINE=89308594; PubMed=2473066;  
 RA Graff J.M., Stumpo D.J., Blackshear P.J.,  
 RT "Characterization of the phosphorylation sites in the chicken and  
 RT bovine myristoylated alanine-rich C kinase substrate protein, a  
 RT prominent cellular substrate for protein kinase C.";  
 RL J. Biol. Chem. 264:11912-11919 (1989).  
 RN [5]  
 RP PHOSPHORYLATION SITES, AND REVISIONS.  
 RC TISSUE=Brain;  
 RX MEDLINE=94308052; PubMed=8034575;  
 RA Taniguchi H., Manenti S., Suzuki M., Titani K.,  
 RT "Myristoylated alanine-rich C kinase substrate (MARCKS), a major  
 RT protein kinase C substrate, is an in vivo substrate of  
 RT proline-directed protein kinase(s). A mass spectroscopic analysis of  
 RT the post-translational modifications.";  
 RL J. Biol. Chem. 269:18299-18302 (1994).  
 RN [6]  
 RP REVERSIBLE ASSOCIATION WITH THE MEMBRANE.  
 RX MEDLINE=91238951; PubMed=2034276;  
 RA Thelen M., Rosen A., Nairn A.C., Aderem A.,  
 RT "Regulation by phosphorylation of reversible association of a  
 RT myristoylated protein kinase C substrate with the plasma membrane.";  
 RL Nature 351:320-322 (1991).  
 RN [7]  
 RP ACTIN-FILAMENT CROSS-LINKING.  
 RX MEDLINE=92220195; PubMed=1560845;  
 RA Hartwig J.H., Thelen M., Rosen A., Janney P.A., Nairn A.C.,  
 RA Aderem A.,  
 RT "MARCKS is an actin filament crosslinking protein regulated by  
 RT protein kinase C and calcium-calmodulin.";  
 RL Nature 356:618-622 (1992).  
 CC -!- FUNCTION: MARCKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR  
 CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND  
 CC SYNAPSIN. MARCKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.  
 CC -!- PHOSPHORYLATION BY PKC DISPLACES MARCKS FROM THE MEMBRANE. IT  
 CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.  
 CC -!- SIMILARITY: Belongs to the MARCKS family.  
 CC -!- CAUTION: REF.1 AND REF.2 SEQUENCES DIFFER FROM THAT SHOWN IN  
 CC POSITIONS 140 TO 150 DUE TO A FRAMESHIFT.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way



DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Inner membrane protein oxalA.  
 GN OXAA OR PRO1302.  
 OS Prochlorococcus marinus.  
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;  
 CC Prochlorococcus.  
 CX NCBI\_TaxID=1219;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SARG / CCMF 1375 / SS120;  
 RX MEDLINE=22810154; PubMed=12917486;  
 RA Dufresne A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,  
 RA Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,  
 RA Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,  
 RA Scanlan D.J., Tandeau de Marsac N., Weissenbach J., Wincker P.,  
 RA Wolf Y.I., Hess W.R.;  
 RT "Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,  
 RT a nearly minimal oxyphototrophic genome.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).  
 CC !- FUNCTION: Required for the insertion of integral membrane proteins  
 CC into the membrane. Probably plays an essential role in the  
 CC integration of proteins of the respiratory chain complexes.  
 CC Involved in integration of membrane proteins that insert  
 CC dependently and independently of the Sec translocase complex (By  
 CC similarity).  
 CC !- SUBUNIT: Specifically interacts with transmembrane segments of  
 CC nascent integral membrane proteins during membrane integration (By  
 CC similarity).  
 CC !- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (By similarity).  
 CC !- SIMILARITY: Belongs to the OXAL/oxalA family. Subfamily 1.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE017164; AAQ00346.1; -  
 CC HAMAP; MF 01810; -; 1.  
 KW Transmembrane; Inner membrane; Complete proteome.  
 FT TRANSMEM 20 42 POTENTIAL.  
 FT TRANSMEM 266 288 POTENTIAL.  
 FT TRANSMEM 303 325 POTENTIAL.  
 SQ SEQUENCE 382 AA; 41425 MW; 12101B3452EB9345 CRC64;  
 Query Match 40.0%; Score 46; DB 1; Length 382;  
 Best Local Similarity 39.1%; Pred. No. 24;  
 Matches 9; Conservative 4; Mismatches 10; Indels 0; Gaps 0;  
 QY 1 GAQFSKTAAGKGAARPGAAV 23  
 DB 210 CADLASVSADGTVTKYFGDATV 232  
 RESULT 8  
 ID VE2 PAPVD STANDARD; PRT; 416 AA.  
 AC P03123;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Probable regulatory protein E2.  
 GN E2.  
 OS Deer papillomavirus.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.  
 CX NCBI\_TaxID=10564;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85293253; PubMed=2993669;

RA Groff D.E., Lancaster W.D.;  
 RT "Molecular cloning and nucleotide sequence of deer papillomavirus.";  
 RL J. Virol. 56:85-91(1985).  
 CC !- FUNCTION: E2 REGULATES VIRAL TRANSCRIPTION AND DNA REPLICATION.  
 CC IT BINDS TO THE E2RE RESPONSE ELEMENT (5'-ACNNNNNGGT-3') PRESENT  
 CC IN MULTIPLE COPIES IN THE REGULATORY REGION. IT CAN EITHER  
 CC ACTIVATE OR REPRESS TRANSCRIPTION DEPENDING ON E2RE'S POSITION  
 CC WITH REGARDS TO PROXIMAL PROMOTER ELEMENTS. REPRESSION OCCURS  
 CC BY STERICALLY HINDERING THE ASSEMBLY OF THE TRANSCRIPTION  
 CC INITIATION COMPLEX. THE E1-E2 COMPLEX BINDS TO THE ORIGIN OF DNA  
 CC REPLICATION.  
 CC !- SUBUNIT: Binds DNA as a dimer (By similarity).  
 CC !- SUBCELLULAR LOCATION: Nuclear.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M11910; AAA66844.1; -  
 CC HSP; P03122; 2BOP.  
 DR InterPro; IPR000427; E2\_C.  
 DR InterPro; IPR001866; E2\_N.  
 DR InterPro; IPR009021; Viral\_DNA\_bd.  
 DR Pfam; PF00511; E2\_C; 1.  
 DR Pfam; PF00508; E2\_N; 1.  
 DR ProDom; PD000672; E2\_C; 1.  
 DR ProDom; PD000678; E2\_N; 1.  
 KW Early protein; Transcription regulation; Activator; DNA-binding;  
 KW Trans-acting factor; DNA replication; Repressor; Nuclear protein.  
 SQ SEQUENCE 416 AA; 46258 MW; 79064F8017368693 CRC64;  
 Query Match 40.0%; Score 46; DB 1; Length 416;  
 Best Local Similarity 60.0%; Pred. No. 26;  
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 9 AKGAAARPGAAV 23  
 DB 368 AVGEQSGRPGDATV 382  
 RESULT 9  
 ID R23B HUMAN STANDARD; PRT; 409 AA.  
 AC P54727; Q8WUB0;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE UV excision repair protein RAD23 homolog B (HHR23B) (XP-C repair  
 DE complementing complex 58 kDa protein) (P58).  
 GN RAD23B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=94222030; PubMed=8168482;  
 RA Masutani C., Sugawara K., Yanagisawa J., Sonoyama T., Ui M.,  
 RA Enomoto T., Takio K., Tanaka K., van der Spek P.J., Bootsma D.,  
 RA Hoeijmakers J.H.J., Hanaoka F.;  
 RT "Purification and cloning of a nucleotide excision repair complex  
 RT involving the Xeroderma pigmentosum group C protein and a human  
 RT homologue of yeast RAD23.";  
 RL EMBO J. 13:1831-1843(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT VAL-249.  
 RA Rieder M.J., Livingston R.J., Daniels M.R., Montoya M.A., Chung M.-W.,  
 RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,  
 RA Schackwitz W.S., Sherwood J.K., Wittrak L.A., Nickerson D.A.;



```

RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ramsay H.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A., AND VARIANT VAL-249.
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Scheinman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Mozny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Touchman J.W., Green E.D., Dickson M.C.,
RA Buterfield V.S.N., Grimwood J., Schmutz J., Myers R.M.,
RA Schnerch A., Schein J.E., Krzywinski M.I., Skalska U., Smalhus D.E.,
RT human and mouse cDNA sequences."
RT "Generation and initial analysis of more than 15,000 full-length
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Involved in DNA excision repair. May play a part in DNA
CC damage recognition and/or in altering chromatin structure to allow
CC access by damage-processing enzymes.
CC -!- SUBUNIT: Heterodimer of a 125 kDa subunit (p125) and of a
CC 58 kDa subunit (p58). Interacts with MJD.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- DOMAIN: The ubiquitin-like domain mediates interaction with MJD.
CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -!- SIMILARITY: Contains 2 UBA domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D21090; BAA04652.1; -
CC EMBL; AL137852; CAD13275.1; -
CC EMBL; AY165178; ANA47194.1; -
CC EMBL; BC020973; AAH20973.1; -
CC EMBL; PIR; S44346; S44346.
CC HSSP; P54725; 1DVO.
CC Genew; HGNC:9813; RAD23B.
CC MIM; 600062; -
CC GO; GO:0005634; C:nucleus; TAS.
CC GO; GO:0003697; F:single-stranded DNA binding; TAS.
CC GO; GO:0006289; P:nucleotide-excision repair; TAS.
CC InterPro; IPR004806; Rad23.
CC InterPro; IPR006636; STI1.
CC InterPro; IPR000449; UBA domain.
CC InterPro; IPR000626; Ubiquitin.
CC Pfam; PF00627; UBA; 2
CC Pfam; PF00240; ubiquitin; 1.
CC SMART; SM00727; STI1; 1.
CC SMART; SM00165; UBA; 2.
CC SMART; SM00213; UBQ; 1.
CC TIGRFAMs; TIGR00601; rad23; 1.
CC PROSITE; PS50030; UBA; 2.
CC PROSITE; PS50033; UBUIQUITIN 2; 1.
KW DNA damage; DNA repair; Nucleic protein; Polymorphism; Repeat.
FT DOMAIN 1 79 UBUIQUITIN-LIKE.
FT DOMAIN 188 228 UBA 1.

```

```

FT DOMAIN 364 404 UBA 2.
FT DOMAIN 103 106 POLY-THR.
FT DOMAIN 254 260 POLY-ALA.
FT DOMAIN 261 269 POLY-THR.
FT DOMAIN 336 348 POLY-GLY.
FT VARIANT 249 249 A -> V (in dbSNP:1805329).
FT SQ SEQUENCE 409 AA; 43171 MW; C026C78273BCB289 CRC64;
Query Match 39.1%; Score 45; DB 1; Length 409;
Best Local Similarity 55.6%; Pred. No. 35;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 7 TAAKGEAAARPPGEEAAYA 24
DB 141 SAAKQEPKPAETTPVA 158
:|||||:|:|:|:|
RESULT 10
PRGR MOUSE STANDARD; PRT; 923 AA.
AC Q00175;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Progesterone receptor (PR).
GN PGR OR NR3C3 OR PR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91299759; PubMed=2069958;
RA Schott D.R., Shyamala G., Schneider W., Parry G.;
RT "Molecular cloning, sequence analyses, and expression of
RT complementary DNA encoding murine progesterone receptor.";
RL Biochemistry 30:7014-7020(1991).
RN [2]
RP SEQUENCE OF 1-9 FROM N.A.
RC STRAIN=129/SV.
RX MEDLINE=95100331; PubMed=7802637;
RA Hagihara K., Wu-Peng X.S., Funabashi T., Kato J., Pfaff D.W.;
RT "Nucleic acid sequence and N-terminus sensitive sites of the 5'
RT region of the mouse progesterone receptor gene.";
RL Biochem. Biophys. Res. Commun. 205:1093-1101(1994).
CC -!- FUNCTION: The steroid hormones and their receptors are involved in
CC the regulation of eukaryotic gene expression and affect cellular
CC proliferation and differentiation in target tissues.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DOMAIN: Composed of three domains: a modulating N-terminal domain,
CC a DNA-binding domain and a C-terminal steroid-binding domain.
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR3
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M68915; AAA39971.1; -
CC EMBL; U12644; AAA66067.1; -
CC PIR; A39596; A39596.
CC HSSP; P06481; 1A28.
CC TRANSFAC; T04680; -.
CC MGD; MGI:97567; Pgr.
CC InterPro; IPR000536; Hormone rec lig.
CC InterPro; IPR00128; Progesterone receptor.
CC InterPro; IPR001723; Steroid receptor.
CC InterPro; IPR008946; Str_ncl_receptor.

```

```
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF02161; Prog_receptor; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRODHOMER.
DR PRINTS; PR00047; STROLDINGER.
DR PRODOM; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Steroid-binding.
FT DOMAIN 1 556 MODULATING, PRO-RICH.
FT DNA_BIND 557 622 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 557 577 C4-TYPE.
FT ZN_FING 593 617 C4-TYPE.
FT DOMAIN 671 923 STEROID-BINDING.
FT DOMAIN 184 188 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 923 AA; 99073 MW; 9415F1ED343BEE3F CRC64;

Query Match 39.1%; Score 45; DB 1; Length 923;
Best Local Similarity 52.4%; Pred. No. 75;
Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 4 FSKTAAGGAAERPGAAVA 24
DB 415 FFLAPAPQAAPSSRPGAAVA 435

RESULT 11
TOP2_PENCH
ID TOP2_PENCH STANDARD; PRT; 1587 AA.
AC Q9Y8G8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE DNA topoisomerase II (EC 5.99.1.3).
GN TOP2.
OS Penicillium chrysogenum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
OX NCBI_TaxID=5076;
RN [1]
SEQUENCE FROM N.A.
RA Kim K., Akashi T., Mizuguchi I., Ozeki M., Kanbe T., Kikuchi A.;
RT "Type II DNA topoisomerase of Penicillium chrysogenum."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Control of topological states of DNA by transient
CC breakage and subsequent rejoining of DNA strands. Topoisomerase II
CC makes double-strand breaks.
CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- MISCELLANEOUS: Eukaryotic topoisomerase I and II can relax both
CC negative and positive supercoils, whereas prokaryotic enzymes
CC relax only negative supercoils.
CC -!- SIMILARITY: Belongs to the type II topoisomerase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB029613; BAA82356.1; -.
CC HSSP; P06786; 1BGW.
DR InterPro; IPR003594; ATPbind ATPase.
DR InterPro; IPR003957; CBPA_NFYB_topis.
DR InterPro; IPR001241; DNA_topoisom.
DR InterPro; IPR002205; DNA_topoisom.

DR Pfam; PF00204; DNA_gyraseB; 1.
DR Pfam; PF00521; DNA_topoisomIV; 1.
DR Pfam; PF02518; HATPase_C; 1.
DR PRINTS; PR00615; CCAATSUBUNTA.
DR PRINTS; PR00418; TPI2FAMILY.
DR PRODOM; PD000742; DNA_topoisomIV; 1.
DR SMART; SM00387; HATPase_C; 1.
DR SMART; SM00433; TOP2C; 1.
DR SMART; SM00434; TOP4C; 1.
DR PROSITE; PS00177; TOPOISOMERASE II; 1.
KW Isomerase; Topoisomerase; DNA-binding; ATP-binding; Phosphorylation;
KW Nuclear protein.
FT NP_BIND 196 201 ATP (POTENTIAL).
FT ACT_SITE 833 833 DNA CLEAVAGE (BY SIMILARITY).
SQ SEQUENCE 1587 AA; 178760 MW; E3A992F63474BD64 CRC64;

Query Match 39.1%; Score 45; DB 1; Length 1587;
Best Local Similarity 57.9%; Pred. No. 1.2e+02;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 6 KTAAGGAAERPGAAVA 24
DB 1524 KPAAGKGAAGKSTAAAA 1542

RESULT 12
USF2_RAT
ID USF2_RAT STANDARD; PRT; 291 AA.
AC Q63665;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Upstream stimulatory factor 2 (Upstream transcription factor 2)
DE (Major late transcription factor 2) (Fragment).
GN USF2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley; TISSUE=Liver;
RX MEDLINE=96139468; PubMed=8576131;
RA Viollet B., Lefrancois-Martinez A.-M., Henrion A., Kahn A.,
RA Raymond-Jean M., Martinez A.;
RT "Immunochemical characterization and transacting properties of
RT upstream stimulatory factor isoforms."
RL J. Biol. Chem. 271:1405-1415(1996).
CC -!- FUNCTION: Transcription factor that binds to a symmetrical DNA
CC sequence (E-boxes) (5'-CAGGTG-3') that is found in a variety of
CC viral and cellular promoters.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Binds DNA as an homodimer or a heterodimer
CC (USF1/USF2).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=1;
CC Comment=At least 2 isoforms are produced;
CC Name=1;
CC IsoId=Q63665-1; Sequence=Displayed;
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X90823; CAA62338.1; -.
CC HSSP; P22415; 1AN4.
DR TRANSFAC; T02115; -.

```

```

DR InterPro: IPR001092; HLH_basic.
DR Pfam: PF00010; HLH; 1.
DR SMART: SM00353; HLH; 1.
DR PROSITE: PS00888; HLH; 1.
KW DNA-binding; Nuclear protein; Transcription regulation;
KW DNA-binding; Nuclear protein; Transcription regulation;
KW Alternative splicing.
FT NON_TER 1
FT DNA_BIND 180 193
FT DOMAIN 194 236
FT DOMAIN 252 273
SQ SEQUENCE 291 AA; 31651 MW; A241C6E9AF6D2424 CRC64;

Query Match 38.3%; Score 44; DB 1; Length 291;
Best Local Similarity 41.7%; Pred. No. 36;
Matches 10; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAGKGAARPGAAVA 24
Db 54 GGQQAQTQGVGDGAQRGPAAAS 77

RESULT 13
USF2 HUMAN STANDARD; PRT; 346 AA.
ID USF2_HUMAN STANDARD; PRT; 346 AA.
AC Q15853; Q00671; Q00709; Q05750; Q07952; Q15851; Q15852;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Upstream stimulatory factor 2 (Upstream transcription factor 2)
DE (FOS-interacting protein) (FIP) (Major late transcription factor 2).
GN USF2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=96139468; PubMed=8576131;
RA Viollet B., Lefrancois-Martinez A.-M., Henrion A., Kahn A.,
RT Raymondjean M., Martinez A.;
RA "Immunological characterization and transacting properties of
RT upstream stimulatory factor isoforms.";
RL J. Biol. Chem. 271:1405-1415(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97127588; PubMed=8954795;
RA Groenen P.M.A., Garcia E., Debeer P., Devriendt K., Frys J.-P.,
RA van de Ven W.J.M.;
RT "Structure, sequence, and chromosome 19 localization of human USF2
RT and its rearrangement in a patient with multicystic renal
RT dysplasia.";
RL Genomics 38:141-148(1996).
RN [3]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Stillwagen S., Ramirez M., Carrano A.;
RT "Characterization by genomic sequence analysis of a gene-rich 1.1 kb
RT region of 19p33.2 containing the human DNA repair gene, RAD23A.";
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM USF2A).
RC TISSUE=Blood;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Tohyiuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.O., Hulyk S.W.,

```

```

RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield F.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 46-346 FROM N.A.
RC TISSUE=B-cell;
RX MEDLINE=93082094; PubMed=1450663;
RA Sirito M., Walker S., Lin Q., Kozlowski M.T., Klein W.H., Sawadogo M.;
RT "Members of the USF family of helix-loop-helix proteins bind DNA as
RT homo- as well as heterodimers.";
RL Gene Expr. 2:231-240(1992).
RN [6]
RP SEQUENCE OF 93-346 FROM N.A.
RX MEDLINE=92271211; PubMed=1589769;
RA Blauar M.A., Rutter W.J.;
RT "Interaction cloning: identification of a helix-loop-helix zipper
RT protein that interacts with c-Fos.";
RL Science 256:1014-1018(1992).
CC -!- FUNCTION: Transcription factor that binds to a symmetrical DNA
CC sequence (E-boxes) (5'-CACGTG-3') that is found in a variety of
CC viral and cellular promoters.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein. Binds DNA as an homodimer or a heterodimer
CC (USF1/USF2). In vivo, the USF1/USF2A heterodimer represents over
CC 66% of the usf binding activity whereas the USF1 and USF2A
CC homodimers represent less than 10%. The USF1/USF2B heterodimer
CC accounted for almost 15% in some cell.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Comment=Additional isoforms seem to exist;
CC Name=USF2A;
CC IsoId=Q15853-1; Sequence=Displayed;
CC Name=USF2A-delta-H;
CC IsoId=Q15853-2; Sequence=VSP_002165;
CC Name=USF2B;
CC IsoId=Q15853-3; Sequence=VSP_002164;
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X90826; CAA62341.1; -
CC EMBL; X90825; CAA62340.1; -
CC EMBL; X90824; CAA62339.1; -
CC EMBL; BC049821; AAH49821.1; -
CC EMBL; Y07661; CAA68942.1; -
CC EMBL; AD000684; AAB51179.1; -
CC EMBL; S50537; AAB24368.1; -
CC EMBL; M77476; -; NOT_ANNOTATED_CDS.
CC HSSP; P22415; 1AN4.
CC TRANSFAC; T00878; -
CC Genew; HGNC:12594; USF2.
CC MIM; 600390;
CC GO; GO:0003702; F:RNA polymerase II transcription factor acti...; TAS.
CC GO; GO:0003700; F:transcription factor activity; TAS.
CC InterPro; IPR001092; HLH_basic.
CC Pfam; PF00010; HLH; 1.
CC SMART; SM00353; HLH; 1.
CC PROSITE; PS00888; HLH; 1.
CC PROSITE; PS00888; HLH; 1.

```

```

KW DNA-binding; Nuclear protein; Transcription regulation;
KW Alternative splicing.
FT DOMAIN 11 20 POLY-ALA.
FT DOMAIN 245 248 POLY-ARG.
FT DNA_BIND 235 248 BASIC DOMAIN.
FT DOMAIN 243 291 HELIX-LOOP-HELIX MOTIF.
FT DOMAIN 307 328 LEUCINE-ZIPPER (POTENTIAL).
FT VARSPLIC 77 143 Missing (in isoform USF2B).
FT VARSPLIC 275 282 /FTId-VSP 002164.
FT VARSPLIC 275 282 Missing (in isoform USF2A-delta-H).
FT CONFLICT 46 64 QTAVALTSVQQAFAFGHNI -> GGTSGGRSGGIQTRVQH
FT CONFLICT 93 100 GDTAGAVS -> EFHSWRRH (IN REF. 6).
FT CONFLICT 122 122 A -> V (IN REF. 6).
SQ SEQUENCE 346 AA; 36955 MW; 78CFE97AC4C10CF CRC64;

Query Match
Best Local Similarity 38.3%; Score 44; DB 1; Length 346;
Matches 10; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGKGAARPGEAAVA 24
Db 109 GQQAVTQGVGDGAQRPGPAAS 132

RESULT 14
ID USF2 MOUSE STANDARD; PRT; 346 AA.
AC Q64705;
DC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Upstream stimulatory factor 2 (Upstream transcription factor 2)
DE (Major late transcription factor 2).
GN USF2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv; TISSUE=Kidney, and Spleen;
RX MEDLINE=95014112; PubMed=7523363;
RA Lin Q., Luo X., Sawadogo M.;
RT "Archaeic structure of the gene encoding transcription factor USF.";
RL J. Biol. Chem. 269:23894-23903 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94173694; PubMed=8127680;
RA Sirito M., Lin Q., Maity T., Sawadogo M.;
RT "Ubiquitous expression of the 43- and 44-kDa forms of transcription factor USF in mammalian cells.";
RL Nucleic Acids Res. 22:427-433 (1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=95293408; PubMed=7774954;
RA Henrion A.A., Martinez A., Mattei M.-G., Kahn A., Raymondjean M.;
RT "Structure, sequence, and chromosomal location of the gene for USF2 transcription factors in mouse.";
RL Genomics 25:36-43 (1995).
CC -!- FUNCTION: Transcription factor that binds to a symmetrical DNA sequence (E-boxes) (5'-CACGTG-3') that is found in a variety of viral and cellular promoters.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another bHLH protein. Binds DNA as an homodimer or a heterodimer (USF1/USF2).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Comment=Additional isoforms seem to exist;
CC Name=USF2A;

```

```

CC Name=USF2B;
CC IsoId=Q64705-1; Sequence=Displayed;
CC Name=USF2B;
CC IsoId=Q64705-2; Sequence=VSP_002166;
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U12283; AAB60674.1; -
DR EMBL; U12282; AAB60674.1; JOINED.
DR EMBL; U01662; AAA20492.1; -
DR EMBL; U01663; AAA20493.1; ALT_INIT.
DR EMBL; X77602; CAA54697.1; -
DR EMBL; X77605; CAA54697.1; JOINED.
DR PIR; A55111; A55111.
DR HSSP; P22415; IAN4.
DR TRANSFAC; T01555; -.
DR MGD; MGI:99961; Usf2.
DR GO; GO:0003677; F:DNA binding; IDA.
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti.; IDA.
DR GO; GO:0006357; P:regulation of transcription from Pol II pro.; IDA.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH; 1.
KW DNA-binding; Nuclear protein; Transcription regulation;
KW Alternative splicing.
FT DOMAIN 14 20 POLY-ALA.
FT DNA_BIND 235 248 BASIC DOMAIN.
FT DOMAIN 249 291 HELIX-LOOP-HELIX MOTIF.
FT DOMAIN 307 328 LEUCINE-ZIPPER (POTENTIAL).
FT VARSPLIC 77 143 Missing (in isoform USF2B).
FT /FTId-VSP 002166.
FT CONFLICT 240 243 MISSING (IN REF. 3).
FT CONFLICT 326 326 A -> S (IN REF. 3).
SQ SEQUENCE 346 AA; 36954 MW; E9D216C25F9447B CRC64;

Query Match
Best Local Similarity 38.3%; Score 44; DB 1; Length 346;
Matches 10; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGKGAARPGEAAVA 24
Db 109 GQQAVTQGVGDGAQRPGPAAS 132

RESULT 15
ID RL3 SCHPO STANDARD; PRT; 110 AA.
AC P17477;
DC 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 60S acidic ribosomal protein p1-alpha 3 (A3).
GN RPA3 OR SPBC3B9.13C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90220620; PubMed=2325655;
RA Beltrame M., Bianchi M.E.;
RT "A gene family for acidic ribosomal proteins in Schizosaccharomycetes
RT pombe: two essential and two nonessential genes.";
RL Mol. Cell. Biol. 10:2341-2348 (1990).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymoprez E.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goifeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Plays an important role in the elongation step of
CC protein synthesis.
CC -!- SUBUNIT: P1 and P2 exist as dimers at the large ribosomal subunit.
CC -!- MISCELLANEOUS: Yeasts contain 4 individual small ribosomal A
CC proteins (RPA) which can be classified into two couples of similar
CC but not identical sequences. Each couple is distinctly related to
CC one of the two A proteins present in multicellular organisms.
CC -!- MISCELLANEOUS: Rpa3 and rpa4 are essential for cell survival,
CC whereas rpa1 and rpa2 are not.
CC -!- SIMILARITY: Belongs to the L12P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M33139; AAA35336.1; -.
CC EMBL; AL022070; CAA17793.1; -.
CC PIR; C34715; R6BYP3.
CC GeneDB.SPombe; SPBC3B9.13c; -.
CC InterPro; IPR001813; Ribosomal_60S.
CC Pfam; PF00428; 60s_ribosomal; 1.
CC Ribosomal protein; Phosphorylation; Multigene family.
CC SK SEQUENCE 110 AA; 11171 MW; 1734AC9779F5A891 CRC64;
QY Query Match 37.4%; Score 43; DB 1; Length 110;
Dy Best Local Similarity 54.5%; Pred. No. 20;
Matches 12; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1 GAQFSKTAAGGAAAEKPEAA 22
Dy 73 GAGAPAAAGGAAAEKPEAA 94

```

Search completed: October 7, 2004, 17:34:16  
Job time : 11 secs

RESULT 1					
ID	O93503	PRELIMINARY;	PRT;	287 AA.	
AC	O93503;				
DT	01-NOV-1998	(TrEMBLrel_08, Created)			
DT	01-NOV-1998	(TrEMBLrel_08, Last sequence update)			
DT	01-OCT-2003	(TrEMBLrel_25, Last annotation update)			
DE	Myristoylated alanine-rich C kinase substrate.				
GN	MARCKS.				
OS	Xenopus laevis (African clawed frog).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;				
OC	Xenopodinae; Xenopus.				
OX	NCBI_TaxID=8355;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RP	MEDLINE=98030614; PubMed=9361009;				
RA	Shi Y., Sullivan S.K., Pitterle D.M., Kennington E.A., Graff J.M.,				
RA	Blackshear P.O.;				
RT	"Mechanisms of MARCKS gene activation during Xenopus development.";				
RL	J. Biol. Chem. 272:29290-29300(1997).				
DR	EMBL; AF017299; AAC61897.1; "				
DR	GO; GO:0005516; F:calmodulin binding; IEA.				
DR	GO; GO:0016301; F:kinase activity; IEA.				
DR	InterPro; IPR002101; MARCKS.				
DR	Fram; PF02063; MARCKS; 1.				
DR	PRINTS; PR00963; MARCKS.				
DR	PROSITE; PS00826; MARCKS_1; 1.				
KW	PROSITE; PS00827; MARCKS_2; 1.				
SQ	SEQUENCE 287 AA; 29147 MW; 35CB7AE609BD3C1 CRC64;				
 Query Match 79.6%; Score 91.5; DB 13; Length 287;					
Best Local Similarity 84.0%; Pred. No. 4.7e-05;					
Matches 21; Conservative 1; Mismatches 2; Indels 1; Gaps					
Qy	1 GAQFSKTAAKGEAA-AERPGEAAVA 24				
Db	2 GAQFSKTAAKGEATAATKPGEAVPA 26				

\*Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

RT Q8AVY4 PRELIMINARY; PRT; 287 AA.  
AC Q8AVY4;  
DT 01-MAR-2003 (TREMBlrel. 23, Created)  
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Embryo;  
RA Klein S., Strausberg R.;  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC041207; AAH41207.1; -  
DR GO; GO:0005516; F:calmodulin binding; IEA.  
DR InterPro; IPR002101; MARCKS.  
DR Pfam; PF02063; MARCKS; 1.  
DR PRINTS; PR00963; MARCKS.  
DR PROSITE; PS00826; MARCKS\_1; 1.  
DR PROSITE; PS00827; MARCKS\_2; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 287 AA; 29064 MW; 90CE6EB6085035FB CRC64;

Query Match 79.6%; Score 91.5; DB 13; Length 287;  
Best Local Similarity 84.0%; Pred. No. 4.7e-05;  
Matches 21; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Qy 1 GAQFSKTAAGKGAAPGAAVA 24  
Db 2 GAQFSKTAAGKGAAPGAAVA 26

RESULT 3  
ID Q7SXW9 PRELIMINARY; PRT; 232 AA.  
AC Q7SXW9;  
DT 01-OCT-2003 (TREMBlrel. 25, Created)  
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Hypothetical protein (Fragment).  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AB; TISSUE=Body;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Udwin T.B., Toehiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;

Query Match 79.6%; Score 91.5; DB 13; Length 287;  
Best Local Similarity 84.0%; Pred. No. 4.7e-05;  
Matches 21; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Qy 1 GAQFSKTAAGKGAAPGAAVA 24  
Db 2 GAQFSKTAAGKGAAPGAAVA 26

RESULT 4  
ID Q9ASIO PRELIMINARY; PRT; 356 AA.  
AC Q9ASIO;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE P0439B06.32 protein (OSJNB0032H19.11 protein).  
GN P0439B06.32.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC  
RT clone:P0439B06.";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, BAC  
RT clone:OSJNB0032H19.";  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP002882; BAB39897.1; -  
DR EMBL; AP003219; BAB61181.1; -  
DR Gramineae; Oryza;  
DR InterPro; IPR008938; ARM.  
DR InterPro; IPR000225; Armadillo.  
DR Pfam; PF00514; Armadillo\_seg; 3.  
DR SMART; SM00185; ARM; 4.  
DR PROSITE; PS0176; ARM\_REPEAT; 1.  
SQ SEQUENCE 356 AA; 38034 MW; 1BC75C2F29197064 CRC64;

Query Match 46.1%; Score 53; DB 10; Length 356;  
Best Local Similarity 75.0%; Pred. No. 19;  
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 TAAGKGAAPGAA 22  
Db 6 TAAGKGAAPGAA 21

RESULT 5  
ID O68475 PRELIMINARY; PRT; 314 AA.  
AC O68475;



DT 01-AUG-1998 (TREMELrel. 07, Created)  
 DT 01-AUG-1998 (TREMELrel. 07, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Chitinase.  
 OS Xanthomonas maltophilia (Pseudomonas maltophilia) (Stenotrophomonas maltophilia).  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;  
 OC Xanthomonadaceae; Stenotrophomonas.  
 OX NCBI\_TaxID=40324;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=kd;  
 RA Kanjanakorn C., Wiwat C.;  
 RT "Chitinase gene from Pseudomonas maltophilia."  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF047411; AAC05490.1; -.  
 DR HSP; P21892; 1EEJ  
 DR InterPro; IPR000437; Prok\_lipoprot.S.  
 DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.  
 SQ SEQUENCE 314 AA; 33421 MW; 8753FBAD9BEC42A8 CRC64;  
  
 Query Match 43.5%; Score 50; DB 2; Length 314;  
 Best Local Similarity 56.5%; Pred. No. 44;  
 Matches 13; Conservative 2; Mismatches 8; Indels 0; Gaps 0;  
  
 QY 2 AQFSKTAAGGAAARPGGAAVA 24  
 || : ||| ||| : ||| |||  
 Db 19 AQPAPPAKAPAAKPGPAANA 41  
  
 RESULT 6  
 Q8U7W4  
 ID Q8U7W4 PRELIMINARY; PRT; 399 AA.  
 AC Q8U7W4;  
 DT 01-JUN-2002 (TREMELrel. 21, Created)  
 DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Hypothetical protein Atu4335.  
 GN ATU4335 OR AGR L 1057.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.  
 OX NCBI\_TaxID=176299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608550; PubMed=11743193;  
 RA Wood D.W., Strubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,  
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
 RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,  
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.;  
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
 C58."  
 RT Science 294:2317-2323 (2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608551; PubMed=11743194;  
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Houmli K., Gordon J., Vaudin M., Iatchouk O., Epp A., Liu F.,  
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
 RA Flanagan C., Crowell C., Gursun J., Lomo C., Sear C., Strub G.,  
 RA Cielo C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 Agrobacterium tumefaciens C58."  
 RT Science 294:2323-2328 (2001).  
 DR EMBL; AE009362; AAL45129.1; -.  
 DR EMBL; AE008251; AAK99105.1; -.

DR PIR; AC3089; AC3089.  
 DR PIR; G98197; G98197.  
 DR InterPro; IPR000253; FHA.  
 DR InterPro; IPR008994; SMAD\_FHA.  
 DR Pfam; PF00498; FHA; 1.  
 DR SMART; SM00240; FHA; 1.  
 DR PROSITE; PS00006; FHA DOMAIN; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 399 AA; 43398 MW; 3719D2DFFEC234E4 CRC64;  
  
 Query Match 43.5%; Score 50; DB 16; Length 399;  
 Best Local Similarity 50.0%; Pred. No. 57;  
 Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
  
 QY 3 QFSKTAAGGAAARPGGAAVA 24  
 || : ||| ||| : ||| |||  
 Db 367 QFLQARGAASARETDATAA 388  
  
 RESULT 7  
 Q8ZYB9  
 ID Q8ZYB9 PRELIMINARY; PRT; 2785 AA.  
 AC Q8ZYB9;  
 DT 01-MAR-2002 (TREMELrel. 20, Created)  
 DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
 DE PARF2b.  
 GN PARF2b.  
 OS Pyrobaculum aerophilum.  
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;  
 OC Thermoproteaceae; Pyrobaculum.  
 OX NCBI\_TaxID=13773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IM2 / ATCC 51768 / DSM 7523;  
 RX MEDLINE=21664397; PubMed=11792869;  
 RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,  
 RA Miller J.H.;  
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum aerophilum."  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:984-989 (2002).  
 RL EMBL; AE009790; AAL63076.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 2785 AA; 300989 MW; 1E144809E1D59C2D CRC64;  
  
 Query Match 42.6%; Score 49; DB 17; Length 2785;  
 Best Local Similarity 57.9%; Pred. No. 5.7e+02;  
 Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
  
 QY 5 SKTAAGGAAARPGGAAV 23  
 || : ||| ||| : ||| |||  
 Db 1343 AELAFKGYAAPPRAEAAAL 1361  
  
 RESULT 8  
 Q8ZH26  
 ID Q8ZH26 PRELIMINARY; PRT; 261 AA.  
 AC Q8ZH26;  
 DT 01-MAR-2002 (TREMELrel. 20, Created)  
 DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
 DE Putative flagellar basal-body rod protein (Flagellar biosynthesis  
 protein, cell-distal portion of basal-body rod).  
 GN FLGG OR YPO0728 OR Y3450.  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Yersinia.  
 OX NCBI\_TaxID=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CO-92 / Biovar Orientalis;  
 RX MEDLINE=21470413; PubMed=11586360;  
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,

```
RA Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagsels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL; AJ414144; CAC89579.1; -
DR EMBL; AE013948; AAM86939.1; -
DR PIR; AH0089; AH0089.
DR DR GO; GO:0009288; C:flagellum (sensu Bacteria); IEA.
DR DR GO; GO:0003774; F:motor activity; IEA.
DR DR GO; GO:0005198; F:structural molecule activity; IEA.
DR DR GO; GO:0001539; P:ciliary/flagellar motility; IEA.
DR DR InterPro; IPR001444; Flag_bb_rod.
DR DR Pfam; PF00460; flg_bb_rod; 1.
DR DR PROSITE; PS00588; FLAGELLA_BB_ROD; 1.
KW Hypothetical protein; Flagellum; Complete proteome.
SQ SEQUENCE 261 AA; 27458 MW; 81F7EDFF28907F5D CRC64;

Query Match 41.7%; Score 48; DB 16; Length 261;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

OY 1 GAQFSKTAAGKGAARPGGAA 22
DB 189 GNLYRETAAGGAEVGVGEEA 210
| : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : :

RESULT 9
ID Q8FSA8 PRELIMINARY; PRT; 287 AA.
AC Q8FSA8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transcription antitermination protein NusG.
CN CE0486.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005215; BAC17296.1; -
DR DR GO; GO:0003711; P:transcriptional elongation regulator activity; IEA.
DR DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR DR InterPro; IPR005924; KOW.
DR DR InterPro; IPR006645; NGN.
DR DR InterPro; IPR001062; NusG.
DR DR InterPro; IPR008991; Transl_sh3_like.
DR DR Pfam; PF00467; KOW; 1.
DR DR Pfam; PF02357; NusG; 1.
DR DR PRINTS; PR00338; NUSGNTSCPFCT.
DR SMART; SM00738; NGN; 1.
```

```
DR TIGRFAMS; TIGR00922; nusG; 1.
DR PROSITE; PS01014; NUSG; 1.
KW Complete proteome.
SQ SEQUENCE 287 AA; 31421 MW; 680A2268D2EC48D4 CRC64;

Query Match 41.7%; Score 48; DB 16; Length 287;
Best Local Similarity 50.0%; Pred. No. 78;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 1 GAQFSKTAAGKGAARPGG 20
DB 28 GETFDEAAPEGEAFAPAE 47
| : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : :

RESULT 10
Q89CP6 PRELIMINARY; PRT; 334 AA.
ID Q89CP6;
AC Q89CP6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BL17750 protein.
GN BL17750.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005963; BAC53015.1; -
DR DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR001107; Band_7.
DR InterPro; IPR001972; Stomatol.
DR Pfam; PF01145; Band_7; 1.
DR DR PRINTS; PR00721; STOMATIN.
DR Complete proteome.
SQ SEQUENCE 334 AA; 36106 MW; 2F7B269526EAB0D8 CRC64;

Query Match 41.7%; Score 48; DB 16; Length 334;
Best Local Similarity 55.6%; Pred. No. 91;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 2 AOFSKTAAGKGAARPG 19
DB 300 ATFGESAASAAAAARRPG 317
| : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : :

RESULT 11
Q81Y95 PRELIMINARY; PRT; 413 AA.
ID Q81Y95;
AC Q81Y95;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serine protease.
GN BA3660.
OS Bacillus anthracis (strain Ames).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=198094;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22608414; PubMed=12721629;
RX Read T.D., Peterson S.N., Tourasse N., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
```

```

Query Match      41.7%; Score 48; DB 10; Length 465;
Best Local Similarity 47.8%; Pred. No. 1.3e+02;
Matches 11; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 2 AOFSTAAKGEEAAERPGEEA 24

```

RESULT 14	
Q9KW22	
ID	PRELIMINARY; PRT; 790 AA.
AC	Q9KW22
DT	01-OCT-2000 (T-EMBLrel. 15, Created)
DT	01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT	01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE	HrpF.
DE	HRPF.
OS	Xanthomonas oryzae (pv. oryzae).
OC	Bacteria, Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC	Xanthomonadaceae; Xanthomonas.
OX	NCBI_TaxID=64187;
[1]	RN
RN	SEQUENCE FROM N.A.
RP	RC STRAIN=NAFF 311018;
RC	MEDLINE=21303248; PubMed=11410350;
PX	Ochiai H., Inoue Y., Hasebe A., Kaku H.;
RA	"Construction and characterization of a Xanthomonas oryzae pv. oryzae
RT	bacterial artificial chromosome library.";
RT	FEMS Microbiol. Lett. 200:59-65(2001).
RL	FEMS Microbiol. Lett. 200:59-65(2001).
RL	EMBL; AB045312; BAB07869.1; --.
DR	InterPro; IPR008718; NoIX.
DR	Fram; PF05819; NoIX; 1.
SQ	SEQUENCE 790 AA; 84883 MW; 73FD1F71106E56B8 CRC64;

Query Match 41.7%; Score 48; DB 2; Length 790;  
Best Local Similarity 59.1%;  
Pred. NO. 2.2e+02;  
Matches 13; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 5 SKTAAK--GEAAARPGRAVA 24  
Db 689 AKGAAGKAGKTAARPSAAFA 710

RESULT 15  
Q9LEZ5 PRELIMINARY; PRT; 1058 AA.  
AC Q9LEZ5;  
DT 01-OCT-2000 (TREMELrel. 15, Created)  
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN T22D6\_50.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
[1]  
RN SEQUENCE FROM N.A.  
RP Bevan M., Terryn N., Ardiles W., Buysshaert C., Dasseville R.,  
RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,  
RA Villarroel R., Gielen J., Van Montagu M., Bancroft I., Mewes H.W.,  
RA Rudd S., Lemcke K., Mayer K.F.X.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
[2]  
RN SEQUENCE FROM N.A.  
RP EU Arabidopsis sequencing project;  
RA Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AL357612; CAB93712.1; -.  
DR FIR; T50496;  
DR GO; GO:000524; F:ATP binding; IEA.  
DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.  
DR GO; GO:0016787; F:Hydrolase activity; IEA.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR InterPro; IPR01410; DEAD.  
DR InterPro; IPR01650; Helicase\_C.  
DR InterPro; IPR00626; Ubiquitin.  
DR Pfam; PF00270; DEAD; 1.  
DR SMART; SM00271; Helicase\_C; 1.  
DR SMART; SM00487; DEXDc; 1.  
DR SMART; SMO0490; HELIC; 1.  
DR PROSITE; PS0053; UBIQUITIN 2; 1.  
KW Hypothetical protein; ATP-binding; Helicase; Hydrolase.  
SQ SEQUENCE 1058 AA; 116856 MW; D95C267BF2672061 CRC64;  
Query Match 41.7%; Score 48; DB 10; Length 1058;  
Best Local Similarity 52.6%; Pred. No. 2.9e+02;  
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 5 SKTAAGGAARPGRAAV 23  
Db 532 SKVSSSEAAADKPSGAIV 550

Search completed: October 7, 2004, 17:35:31  
Job time : 38.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:32:02 ; Search time 50.5 Seconds  
(without alignments)  
134.280 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAGKGAARPGAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	24	3	AAY95896 MANS pept
2	115	100.0	24	6	ABP97765 Myristoyl
3	115	100.0	165	7	ABP97765 Microsate
4	115	100.0	182	7	ABP97765 Microsate
5	115	100.0	308	7	ABP97765 Microsate
6	115	100.0	308	7	ABP97765 Microsate
7	115	100.0	308	7	ABP97765 Microsate
8	115	100.0	308	7	ABP97765 Microsate
9	115	100.0	330	2	AA05528 High dens
10	115	100.0	332	3	AA05528 Human myr
11	115	100.0	332	3	AA05528 Human myr
12	115	100.0	332	7	ABP97765 Microsate
13	111	96.5	309	7	ABP97765 Microsate
14	49	42.6	380	4	ABU18252 Protein e
15	48	41.7	413	6	ABU18252 Protein e
16	47	40.9	104	7	ABU18252 Protein e
17	47	40.9	341	4	ABU18252 Protein e
18	47	40.9	410	6	ABU18252 Protein e
19	47	40.9	559	6	ABU18252 Protein e
20	46.5	40.4	1128	3	AA05528 Human myr
21	46.5	40.4	1206	3	AA05528 Human myr
22	46.5	40.4	1293	3	AA05528 Human myr
23	46	40.0	916	4	ABP97765 Microsate
24	45	39.1	89	4	AAW21657 Peptide #
25	45	39.1	89	4	ABB44016 Peptide #

#### ALIGNMENTS

##### RESULT 1

AA95896  
ID AAY95896 standard; peptide; 24 AA.

XX AC AAY95896;

XX 20-NOV-2000 (first entry)

DE MANS peptide inhibitor of MARCKS-related mucus secretion.

XX MANS; MARCKS; myristoylated alanine-rich C kinase substrate; human;  
mucus secretion; inhibitor; bronchitis; asthma; cystic fibrosis;  
chronic obstructive pulmonary disease; pneumonia; emphysema; influenza;  
rhinitis; therapy.

XX Homo sapiens.

XX WO200050062-A2.

XX 31-AUG-2000.

XX 24-FEB-2000; 2000WO-US005050.

XX 24-FEB-1999; 99US-00256154.

XX (UYN-) UNIV NORTH CAROLINA STATE.

XX Li Y, Martin LD, Adler KB;

XX WPI; 2000-572036/53.

Regulating mucus secretion by a mucus-secreting cell, useful for treating e.g. bronchitis, asthma or pneumonia, by administering a compound that inhibits or enhances myristoylated alanine-rich C-kinase substrate protein.

XX Claim 12; Page 40; 66pp; English.

CC The present sequence is that of MANS peptide, comprising the N-terminal region of human myristoylated alanine-rich C kinase substrate MARCKS protein (see AAY95896), a major cellular substrate for protein kinase S. MANS peptide inhibits secretion of mucus from mucus membranes and mucus-secreting cells, including human airway epithelial cells. It is suggested to block attachment of MARCKS protein to the mucin granule, thus blocking or inhibiting the release of mucin granules and the secretion of mucus by the cell. The invention relates to methods and compounds for decreasing mucus secretion, particularly in the airways. Such compounds include MANS peptide and antisense oligonucleotides to MARCKS. They are useful in

26	45	39.1	89	4	AA037960	Peptide #
27	45	39.1	89	4	ABP26922	Protein #
28	45	39.1	89	4	AA077746	Human bra
29	45	39.1	89	4	AA065031	Human bra
30	45	39.1	89	4	ABG59403	Human liv
31	45	39.1	89	5	ABG46774	Human pep
32	45	39.1	121	5	ABP30384	Streptoco
33	45	39.1	126	5	ABP28368	Streptoco
34	45	39.1	165	3	AA051017	Arabidops
35	45	39.1	210	3	AA051016	Arabidops
36	45	39.1	263	3	AA051015	Arabidops
37	45	39.1	347	3	AA051011	Arabidops
38	45	39.1	355	3	AA028692	Arabidops
39	45	39.1	392	3	AA051010	Arabidops
40	45	39.1	400	3	AA028691	Arabidops
41	45	39.1	405	3	AA015702	Arabidops
42	45	39.1	405	3	AA054040	Arabidops
43	45	39.1	409	2	AA075700	Vpr prote
44	45	39.1	409	2	AA068186	Vpr bindi
45	45	39.1	409	6	ABU07460	Protein d

CC inhibiting mucus secretion in conditions such as bronchitis, cystic  
CC fibrosis, chronic obstructive pulmonary disease, asthma, emphysema,  
CC pneumonia, influenza, rhinitis and the common cold  
XX Sequence 24 AA;  
SQ Query Match 100.0%; Score 115; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 9.6e-11;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAQFSKTAAGGAAERPGGAAVA 24  
DB 1 GAQFSKTAAGGAAERPGGAAVA 24  
RESULT 2  
ABP97765  
ID ABP97765 standard; peptide; 24 AA.  
XX AC ABP97765;  
XX 11-AUG-2003 (first entry)  
XX Myristoylated N-terminal sequence (MANS) peptide.  
XX Myristoylated N-terminal sequence; MANS; MARCKS; mucus secretion;  
XX inflammatory mediator; inflammation; respiratory disease; asthma;  
XX chronic bronchitis; chronic obstructive pulmonary disease; COPD;  
XX bowel disease; irritable bowel syndrome; Crohn's disease;  
XX ulcerative colitis; skin disease; rosacea; eczema; psoriasis; acne;  
XX autoimmune disease; pain; arthritis; cystic fibrosis.  
XX Synthetic.  
XX Key Location/Qualifiers  
XX Modified-site 1 /note= "N-terminal myristate chain attached"  
XX WO2003000027-A2.  
XX 03-JAN-2003.  
XX 26-JUN-2002; 2002WO-US022270.  
XX 26-JUN-2001; 2001US-0300933P.  
XX (UYNC-) UNIV NORTH CAROLINA STATE.  
XX Martin LD, Adler KB, Li Y;  
XX WPI; 2003-278239/27.  
XX Method of regulating inflammation comprises administering a composition  
XX comprising a MANS peptide or an active fragment thereof.  
XX Disclosure; Page 26; 54pp; English.  
XX The present sequence represents a myristoylated N-terminal sequence  
XX (MANS) peptide. The MANS peptide is identical to the first 24 amino acids  
XX of MARCKS, and mediates insertion of MARCKS into membranes. The MANS  
XX peptide inhibits both mucus secretion and inflammatory mediators. The  
XX specification describes a method of regulating inflammation. The method  
XX comprises administering a composition comprising a MANS peptide. The  
XX method is useful for treating inflammation caused by respiratory diseases  
XX (e.g. asthma, chronic bronchitis and chronic obstructive pulmonary  
XX disease (COPD), bowel diseases (e.g. irritable bowel syndrome, Crohn's  
XX disease and ulcerative colitis), skin diseases (e.g. rosacea, eczema,  
XX psoriasis and severe acne), autoimmune diseases and pain syndromes,  
XX arthritis and cystic fibrosis  
SQ Sequence 24 AA;  
Query Match 100.0%; Score 115; DB 6; Length 24;

Best Local Similarity 100.0%; Pred. No. 9.6e-11;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAQFSKTAAGGAAERPGGAAVA 24  
DB 1 GAQFSKTAAGGAAERPGGAAVA 24  
RESULT 3  
ADE80715  
ID ADE80715 standard; protein; 165 AA.  
XX AC ADE80715;  
XX 29-JAN-2004 (first entry)  
XX Microsatellite related MACS (-1) ORF amino acid sequence.  
XX frameshift mutation; microsatellite; cytostatic; neuroprotective;  
XX vasotropic; vaccine; gene therapy; neurodegenerative disorder;  
XX vascular disease; cancer.  
XX Unidentified.  
XX WO2003087162-A2.  
XX 23-OCT-2003.  
XX 17-APR-2003; 2003WO-EP004083.  
XX 18-APR-2002; 2002EP-00008771.  
XX 18-APR-2002; 2002EP-00008773.  
XX 18-APR-2002; 2002EP-00008774.  
XX (MTM-) MTM LAB AG.  
XX Von Knebel Doebritz M, Gebert J, Linnebacher M, Woerner S;  
XX Ridder R, Bork P, Yuan YP;  
XX WPI; 2003-845308/78.  
XX New nucleic acid, useful in preparing a composition for diagnosing or  
XX treating disorders associated with frameshift mutations in coding  
XX microsatellite regions, e.g., neurodegenerative disorder, vascular  
XX disease or cancer.  
XX Claim 3; Fig 2; 62pp; English.  
XX The present invention describes a nucleic acid sequence (I) which encodes  
XX a polypeptide consisting of TAF1B, MACS, UVRAG, ELAVL3, TCF6L1, ABCF1,  
XX AIM2, CHD2, FL J11053, KIAA1052, ACVR2 or HT001 having a frameshift  
XX mutation. Also described: (1) a frameshift polypeptide (II); (2) a method  
XX for treating disorders associated with frameshift mutations in coding  
XX microsatellites; (3) a pharmaceutical composition comprising the nucleic  
XX acid and/or polypeptide; (4) a method for detecting a disorder associated  
XX with frameshift mutations in coding microsatellite regions; (5) a  
XX diagnostic or research kit for detecting a disorder associated with  
XX frameshift mutations in coding microsatellite regions, comprising the  
XX nucleic acid and/or frameshift polypeptide; and (6) a method for treating  
XX disorders associated with peptides arising from frameshift mutations in  
XX coding microsatellite regions in individuals. (I) and (II) have  
XX cytostatic, neuroprotective and vasotropic activities, and can be used in  
XX vaccines and in gene therapy. The nucleic acid (I) or frameshift  
XX polypeptide (II) can be used in detecting disorders associated with  
XX frameshift mutations in coding microsatellite regions or in preparing  
XX pharmaceutical compositions for treating disorders associated with  
XX frameshift mutations in coding microsatellite regions, e.g.,  
XX neurodegenerative disorder, vascular disease, cancer or precursory stages  
XX of cancer. The present sequence represents a protein used in the  
XX exemplification of the present invention.  
SQ Sequence 165 AA;

Sequence 182 AA;



CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 308 AA;

Query Match 100.0%; Score 115; DB 7; Length 308;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAQFSKTAAGGAAERPGGAAVA 24  
|||||  
Db 1 GAQFSKTAAGGAAERPGGAAVA 24

RESULT 6  
ADE60423  
ID ADE60423 standard; protein; 308 AA.  
XX  
AC ADE60423;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Rat Protein P30009, SEQ ID NO 6332.  
XX  
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX  
OS Rattus norvegicus.  
XX  
FN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
XX (GEHO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
XX Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
DR GENBANK; P30009.  
XX  
XX New composition comprising two or more isolated polypeptides, useful for  
XX preparing a medicament for treating pain in an animal.  
XX  
XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 308 AA;

Query Match 100.0%; Score 115; DB 7; Length 308;  
Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAQFSKTAAGGAAERPGGAAVA 24  
|||||  
Db 1 GAQFSKTAAGGAAERPGGAAVA 24

RESULT 7  
ADE60493  
ID ADE60493 standard; protein; 308 AA.  
XX  
AC ADE60493;  
XX  
DT 29-JAN-2004 (first entry)  
XX  
DE Rat Protein P30009, SEQ ID NO 6402.  
XX  
XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
XX chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX  
OS Rattus norvegicus.  
XX  
FN WO2003016475-A2.  
XX  
PD 27-FEB-2003.  
XX  
PF 14-AUG-2002; 2002WO-US025765.  
XX  
PR 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
XX (GEHO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
XX Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
DR GENBANK; P30009.  
XX  
XX New composition comprising two or more isolated polypeptides, useful for  
XX preparing a medicament for treating pain in an animal.

Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating

CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 308 AA;

Query Match 100.0%; Score 115; DB 7; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAERPGGAAVA 24  
 |||||  
 Db 1 GAQFSKTAAGGAAERPGGAAVA 24

RESULT 8

ADD45848  
 ID ADD45848 standard; protein; 308 AA.

AC ADD45848;

XX 29-JAN-2004 (first entry)

XX Rat Protein P30009, SEQ ID NO 11518.

DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GSHO ) GEN HOSPITAL CORP.

PA (FARB ) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENEANK; P30009.

XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the

CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 308 AA;

Query Match 100.0%; Score 115; DB 7; Length 308;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-09;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAERPGGAAVA 24  
 |||||  
 Db 1 GAQFSKTAAGGAAERPGGAAVA 24

RESULT 9

AAR05528

ID AAR05528 standard; protein; 330 AA.

XX AAR05528;

XX 25-MAR-2003 (revised)

DT 23-OCT-1990 (first entry)

XX High density lipoprotein (HDL) binding protein.

XX High density lipoprotein; HDL-binding protein; atherosclerosis;  
 KW hypercholesterolaemia; ds.

XX Homo sapiens.

XX WO9005744-A.

XX 31-MAY-1990.

XX 18-NOV-1988; 88US-00273388.

XX 18-NOV-1988; 88US-00273388.

XX 05-MAY-1989; 89US-00347855.

XX (UNIW ) UNIV WASHINGTON.

PA (ZYMO ) ZYMOGENETICS INC.

XX Oram JE, McKnight GL, Hart CE, Curtis DA;

XX WPI; 1990-193405/25.

XX N-PSDB; AAQ04784.

XX New mammalian proteins binding high density lipoprotein sub-class 3 - DNA  
 PT encoding them and derived antibodies, for screening potentially  
 PT therapeutic HDL analogues and for diagnosing risk of atherosclerosis.

XX Claim 4; Fig 1A-D; 79pp; English.

XX The protein product may be used to raise Abs, and the cDNA to create  
 CC probes, both useful in screening for HDL analogues, agonists and  
 CC antagonists, and in identifying abnormalities in the HDL binding/receptor  
 CC pathway. HDL analogues can be used in treating hypercholesterolaemia and  
 CC atherosclerosis (Updated on 25-MAR-2003 to correct PR field.) (Updated on  
 CC 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI  
 CC field.)

```
SQ Sequence 330 AA;
Query Match 100.0%; Score 115; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAQFSKTAAGKGEAAERPGGAAVA 24
   |||||
Db 2 GAQFSKTAAGKGEAAERPGGAAVA 25
   |||||

RESULT 10
AAY95898
ID AAY95898 standard; protein; 332 AA.
XX
AC AAY95898;
XX
DT 20-NOV-2000 (first entry)
XX
DE Human myristoylated alanine-rich C kinase substrate MARCKS.
XX
KW MARCKS; myristoylated alanine-rich C kinase substrate; human;
KW mucus secretion; inhibitor; bronchitis; asthma; cystic fibrosis;
KW chronic obstructive pulmonary disease; pneumonia; emphysema; influenza;
KW rhinitis; therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 2..25
FT /note= "MANS peptide of AAY95898"
FT Misc-difference 84
FT /note= "Ser in sequence of AAY95899"
FT Misc-difference 119
FT /note= "Ala in sequence of AAY95899"
FT Peptide 152..176
FT /note= "MA-PSD peptide of AAY95897"
XX
PN WO200050062-A2.
XX
PD 31-AUG-2000.
XX
PF 24-FEB-2000; 2000WO-US005050.
XX
PR 24-FEB-1999; 99US-00256154.
XX
PA (UYN-) UNIV NORTH CAROLINA STATE.
XX
PI Li Y, Martin LD, Adler KB;
XX
PI WPI; 2000-572036/53.
XX
DR N-PSDB; AAA50339.
XX
PT Regulating mucus secretion by a mucus-secreting cell, useful for treating
PT e.g. bronchitis, asthma or pneumonia, by administering a compound that
PT inhibits or enhances myristoylated alanine-rich C-kinase substrate
PT protein.
XX
PS Claim 3; Page 42-43; 66pp; English.
XX
CC The present sequence is that of human myristoylated alanine-rich C kinase
CC substrate MARCKS protein, a major cellular substrate. The invention
CC relates to methods of inhibiting mucus secretion by a mucus-secreting
CC cell by administering a compound that inhibits MARCKS protein-related
CC mucus secretion. Such compounds include active fragments of MARCKS
CC protein such as MANS peptide (see AAY95897) and MA-PSD peptide (see
CC AAY95897), which corresponds to a phosphorylation site of MARCKS. The
CC inhibitor compounds can be used to treat conditions such as bronchitis,
CC cystic fibrosis, chronic obstructive pulmonary disease, asthma,
CC emphysema, pneumonia, influenza, rhinitis and the common cold. An
CC alternative sequence for MARCKS is provided in AAY95899, which differs
CC from the present sequence at 2 amino acid residues, Ala-84 (Ser) and Pro-
CC 119 (Ala)
```

```
XX
SQ Sequence 332 AA;
Query Match 100.0%; Score 115; DB 3; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAQFSKTAAGKGEAAERPGGAAVA 24
   |||||
Db 2 GAQFSKTAAGKGEAAERPGGAAVA 25
   |||||

RESULT 11
AAY95899
ID AAY95899 standard; protein; 332 AA.
XX
AC AAY95899;
XX
DT 20-NOV-2000 (first entry)
XX
DE Human myristoylated alanine-rich C kinase substrate MARCKS.
XX
KW MARCKS; myristoylated alanine-rich C kinase substrate; human;
KW mucus secretion; inhibitor; bronchitis; asthma; cystic fibrosis;
KW chronic obstructive pulmonary disease; pneumonia; emphysema; influenza;
KW rhinitis; therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 2..25
FT /note= "MANS peptide of AAY95896"
FT Misc-difference 84
FT /note= "Ala in sequence of AAY95898"
FT Misc-difference 119
FT /note= "Pro in sequence of AAY95898"
FT Peptide 152..176
FT /note= "MA-PSD peptide of AAY95897"
XX
PN WO200050062-A2.
XX
PD 31-AUG-2000.
XX
PF 24-FEB-2000; 2000WO-US005050.
XX
PR 24-FEB-1999; 99US-00256154.
XX
PA (UYN-) UNIV NORTH CAROLINA STATE.
XX
PI Li Y, Martin LD, Adler KB;
XX
PI WPI; 2000-572036/53.
XX
DR N-PSDB; AAA50340.
XX
PT Regulating mucus secretion by a mucus-secreting cell, useful for treating
PT e.g. bronchitis, asthma or pneumonia, by administering a compound that
PT inhibits or enhances myristoylated alanine-rich C-kinase substrate
PT protein.
XX
PS Disclosure; Page 46-47; 66pp; English.
XX
CC The present sequence is that of human myristoylated alanine-rich C kinase
CC substrate MARCKS protein, a major cellular substrate. The invention
CC relates to methods of inhibiting mucus secretion by a mucus-secreting
CC cell by administering a compound that inhibits MARCKS protein-related
CC mucus secretion. Such compounds include active fragments of MARCKS
CC protein such as MANS peptide (see AAY95897) and MA-PSD peptide (see
CC AAY95897), which corresponds to a phosphorylation site of MARCKS. The
CC inhibitor compounds can be used to treat conditions such as bronchitis,
CC cystic fibrosis, chronic obstructive pulmonary disease, asthma,
CC emphysema, pneumonia, influenza, rhinitis and the common cold. An
CC alternative sequence for MARCKS is provided in AAY95898, which differs
CC from the present sequence at 2 amino acid residues, Ser-84 (Ala) and Ala-
```

CC 119 (Pro)  
XX Sequence 332 AA;  
SQ Sequence 332 AA;  
Query Match 100.0%; Score 115; DB 3; Length 332;  
Best Local Similarity 100.0%; Pred. No. 1.6e-09;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAQFSKTAAGGAAARPGAAVA 24  
|||  
Db 2 GAQFSKTAAGGAAARPGAAVA 25  
|||  
RESULT 12  
ID ADE80714 standard; protein; 332 AA.  
XX ADE80714;  
XX 29-JAN-2004 (first entry)  
XX Microsatellite related MACS wt ORF amino acid sequence.  
XX frameshift mutation; microsatellite; cytostatic; neuroprotective;  
KW vasotropic; vaccine; gene therapy; neurodegenerative disorder;  
KW vascular disease; cancer.  
XX Unidentified.  
XX WO2003087162-A2.  
XX 23-OCT-2003.  
XX 17-APR-2003; 2003WO-EP004083.  
XX 18-APR-2002; 2002EP-0008771.  
PR 18-APR-2002; 2002EP-0008773.  
XX 18-APR-2002; 2002EP-0008774.  
XX (MTM-) MTM LAB AG.  
XX Von Knebel Doeberitz M, Gebert J, Linnebacher M, Woerner S;  
PI Ridder R, Bork P, Yuan YP;  
XX WPI; 2003-845308/78.  
XX New nucleic acid, useful in preparing a composition for diagnosing or  
PT treating disorders associated with frameshift mutations in coding  
PT microsatellite regions, e.g., neurodegenerative disorder, vascular  
PT disease or cancer.  
XX Claim 3; Fig 2; 62pp; English.  
XX The present invention describes a nucleic acid sequence (I) which encodes  
CC a polypeptide consisting of TAF1B, MACS, UVRAG, ELAVL3, TCF6L1, ABCF1,  
CC AIM2, CHD2, FL J11053, KIAA1052, ACVR2 or HT001 having a frameshift  
CC mutation. Also described: (1) a frameshift polypeptide (II); (2) a method  
CC for treating disorders associated with frameshift mutations in coding  
CC microsatellites; (3) a pharmaceutical composition comprising the nucleic  
CC acid and/or polypeptide; (4) a method for detecting a disorder associated  
CC with frameshift mutations in coding microsatellite regions; (5) a  
CC diagnostic or research kit for detecting a disorder associated with  
CC frameshift mutations in coding microsatellite regions, comprising the  
CC nucleic acid and/or frameshift polypeptide; and (6) a method for treating  
CC disorders associated with peptides arising from frameshift mutations in  
CC coding microsatellite regions in individuals. (I) and (II) have  
CC cytostatic, neuroprotective and vasotropic activities, and can be used in  
CC vaccines and in gene therapy. The nucleic acid (I) or frameshift  
CC polypeptide (II) can be used in detecting disorders associated with  
CC frameshift mutations in coding microsatellite regions or in preparing  
CC pharmaceutical compositions for treating disorders associated with  
CC frameshift mutations in coding microsatellite regions, e.g.,  
CC neurodegenerative disorder, vascular disease, cancer or precursory stages

CC of cancer. The present sequence represents a protein used in the  
XX exemplification of the present invention.  
XX  
SQ Sequence 332 AA;  
Query Match 100.0%; Score 115; DB 7; Length 332;  
Best Local Similarity 100.0%; Pred. No. 1.6e-09;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAQFSKTAAGGAAARPGAAVA 24  
|||  
Db 2 GAQFSKTAAGGAAARPGAAVA 25  
|||  
RESULT 13  
ADD47655  
ID ADD47655 standard; protein; 309 AA.  
XX ADD47655;  
XX 29-JAN-2004 (first entry)  
XX Rat Protein NP\_032564, SEQ ID NO 13351.  
DE Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;  
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.  
XX Rattus norvegicus.  
XX WO2003016475-A2.  
XX 27-FEB-2003.  
XX 14-AUG-2002; 2002WO-US025765.  
XX 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
XX 26-NOV-2001; 2001US-0333347P.  
XX (GEHO) GEN HOSPITAL CORP.  
PA (FARB) BAYER AG.  
XX Woolf C, D'urso D, Befort K, Costigan M;  
PI WPI; 2003-268312/26.  
XX GENBANK; NP\_032564.  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
PT  
XX Claim 1; Page; 1017pp; English.  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are: (1) a vector comprising a promoter operably linked to the nucleic acid encoding a polypeptide whose expression is inhibited by the antisense nucleic acid; (2) a host cell containing the vector; (3) an isolated polypeptide or its fragment whose expression is inhibited by the antisense nucleic acid; (4) an antibody capable of specifically binding the polypeptide; (5) producing a polypeptide; (6) inhibiting cellular proliferation or the activity of a gene in an operon required for proliferation; (7) identifying a compound that influences the activity of the gene product or that has an activity against a biological pathway required for proliferation, or that inhibits cellular proliferation; (8) identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a

Sequence 413 AA;

QY      1 GAQFSKTAAKGEAAAERPGEAAVA 24  
         ||| :||| ::||| :|||  
Db    181 GANVNKVATLGDSSKIRAGEKAIA 204

Search completed: October 7, 2004, 17:33:56  
Job time : 56.5 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:35:42 ; Search time 84.5 Seconds  
(without alignments)  
91.399 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAGKGEAAERPGEAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

```

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	100.0	24	US-10-180-753-1	Sequence 1, Appli
2	115	100.0	24	US-10-802-644-1	Sequence 1, Appli
3	115	100.0	389	US-10-131-410-172	Sequence 172, App
4	53	46.1	356	US-10-437-963-185520	Sequence 185520,
5	51	44.3	218	US-10-767-701-46204	Sequence 46204, A
6	51	44.3	1152	US-10-437-963-131104	Sequence 131104,
7	49.5	43.0	254	US-10-437-963-204940	Sequence 204940,
8	48.5	42.2	549	US-10-369-493-10346	Sequence 10346, A
9	48	41.7	148	US-10-767-701-37911	Sequence 37911, A
10	48	41.7	465	US-10-282-122A-46176	Sequence 46176, A
11	48	41.7	981	US-10-437-963-171398	Sequence 171398,
12	48	41.7	981	US-10-437-963-171398	Sequence 171398,
13	47.5	41.3	467	US-10-424-599-169015	Sequence 169015,
14	47	40.9	104	US-10-393-449-41	Sequence 41, Appl
15	47	40.9	104	US-10-177-725-41	Sequence 41, Appl

```

16 47 40.9 176 16 US-10-767-701-56335 Sequence 56335, A
17 47 40.9 200 16 US-10-767-701-35982 Sequence 35982, A
18 47 40.9 410 12 US-10-282-122A-46535 Sequence 46535, A
19 47 40.9 559 12 US-10-282-122A-69664 Sequence 69664, A
20 47 40.9 624 16 US-10-437-963-149882 Sequence 149882,
21 47 40.9 1350 16 US-10-647-196-27 Sequence 27, Appl
22 46.5 40.4 66 16 US-10-437-963-113367 Sequence 113367,
23 46 40.0 60 16 US-10-437-963-172410 Sequence 172410,
24 46 40.0 110 16 US-10-437-963-169526 Sequence 169526,
25 46 40.0 124 16 US-10-437-963-136066 Sequence 136066,
26 46 40.0 125 12 US-10-424-599-227273 Sequence 227273,
27 46 40.0 126 16 US-10-437-963-161220 Sequence 161220,
28 46 40.0 191 16 US-10-437-963-105384 Sequence 105384,
29 46 40.0 217 12 US-10-425-114-49555 Sequence 49555, A
30 46 40.0 324 16 US-10-437-963-144120 Sequence 144120,
31 46 40.0 439 16 US-10-437-963-188362 Sequence 188362,
32 46 40.0 612 16 US-10-437-963-185375 Sequence 185375,
33 45 39.1 89 9 US-09-864-761-42220 Sequence 42220, A
34 45 39.1 148 14 US-10-156-761-8289 Sequence 8289, Ap
35 45 39.1 239 16 US-10-437-963-191253 Sequence 191253,
36 45 39.1 255 16 US-10-437-963-204931 Sequence 204931,
37 45 39.1 469 16 US-10-437-963-126692 Sequence 126692,
38 45 39.1 1004 16 US-10-437-963-109426 Sequence 109426,
39 44.5 38.7 132 16 US-10-437-963-125382 Sequence 125382,
40 44.5 38.7 413 14 US-10-156-761-12219 Sequence 12219, A
41 44.5 38.7 440 15 US-10-369-493-8931 Sequence 8931, Ap
42 44 38.3 98 9 US-09-216-393-98 Sequence 98, Appl
43 44 38.3 98 14 US-10-321-856-98 Sequence 98, Appl
44 44 38.3 99 12 US-10-425-114-37338 Sequence 37338, A
45 44 38.3 111 16 US-10-767-701-53396 Sequence 53396, A

```

#### ALIGNMENTS

```

RESULT 1
US-10-180-753-1
; Sequence 1, Application US/10180753
; Publication No. US20030013652A1
; GENERAL INFORMATION:
; APPLICANT: Martin, Linda
; APPLICANT: Adler, Kenneth
; APPLICANT: Li, Yuehua
; TITLE OF INVENTION: BLOCKING PEPTIDE FOR INFLAMMATORY CELL SECRETION
; FILE REFERENCE: 5051.574
; CURRENT APPLICATION NUMBER: US/10/180.753
; CURRENT FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: N-terminal myristoylation
US-10-180-753-1

```

```

Query Match 100.0%; Score 115; DB 12; Length 24;
Best Local Similarity 100.0%; Pred. No. 8.6e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 GAQFSKTAAGKGEAAERPGEAAVA 24
|||||
Db 1 GAQFSKTAAGKGEAAERPGEAAVA 24
|||||

```

```

RESULT 2
US-10-802-644-1
; Sequence 1, Application US/10802644

```

```
; Publication No. US20040180836A1
; GENERAL INFORMATION:
; APPLICANT: Martin, Linda
; APPLICANT: Adler, Kenneth
; APPLICANT: Li, Yuehua
; TITLE OF INVENTION: BLOCKING PEPTIDE FOR INFLAMMATORY CELL SECRETION
; FILE REFERENCE: 5051,574
; CURRENT APPLICATION NUMBER: US/10/802,644
; CURRENT FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: US/10/180,753
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: N-terminal myristoylation
US-10-802-644-1

Query Match          100.0%; Score 115; DB 16; Length 24;
Best Local Similarity 100.0%; Pred. No. 8.6e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAGGAAARPGAAVA 24
Db 1 GAQFSKTAAGGAAARPGAAVA 24

RESULT 3
US-10-131-410-172
; Sequence 172, Application US/10131410
; Publication No. US20030235915A1
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
; FILE REFERENCE: SCH-1763
; CURRENT APPLICATION NUMBER: US/10/131,410
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 09/646,673
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: PCT/DE99/00908
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 172
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-131-410-172

Query Match          100.0%; Score 115; DB 15; Length 389;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAGGAAARPGAAVA 24
Db 59 GAQFSKTAAGGAAARPGAAVA 82

RESULT 4
US-10-437-963-185520
```

```
; Sequence 185520, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 185520
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_82408C.1.pep
US-10-437-963-185520

Query Match          46.1%; Score 53; DB 16; Length 356;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 TAAKGEAAARPGEEAA 22
Db 6 TAAAGGAAARPGSEAA 21
```

```
RESULT 5
US-10-767-701-46204
; Sequence 46204, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 46204
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C572_1.pep
US-10-767-701-46204

Query Match          44.3%; Score 51; DB 16; Length 218;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 6 KTAAGGAAARPGEEAAV 23
Db 33 RASAEAGGAAAPGRRV 50
```

```
RESULT 6
US-10-437-963-131104
; Sequence 131104, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```

## RESULT 8

```

1 / APPLICANT: Cao Yongwei
2 / TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
3 / TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
4 / FILE REFERENCE: 38-21(53223)B
5 / CURRENT APPLICATION NUMBER: US/10/424,599
6 / CURRENT FILING DATE: 2003-04-28
7 / NUMBER OF SEQ ID NOS: 285684
8 / SEQ ID NO 169015
9 / LENGTH: 467
10 / TYPE: PRT
11 / ORGANISM: Glycine max
12 / FEATURE:
13 / OTHER INFORMATION: Clone ID: PAT MRT3847 123635C.1.pep

```



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:32:03 ; Search time 15 seconds  
(without alignments)  
82.602 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAGKGEAAERPGEAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/2/iaa/5A COMB.pep:\*
- 2: /cgn2\_6/prodata/2/iaa/5B COMB.pep:\*
- 3: /cgn2\_6/prodata/2/iaa/6A COMB.pep:\*
- 4: /cgn2\_6/prodata/2/iaa/6B COMB.pep:\*
- 5: /cgn2\_6/prodata/2/iaa/PCTUS COMB.pep:\*
- 6: /cgn2\_6/prodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	100.0	309	2	US-08-405-175A-8
2	115	100.0	332	2	US-08-405-175A-5
3	111	96.5	309	2	US-08-405-175A-7
4	111	96.5	335	2	US-08-405-175A-6
5	104	90.4	281	2	US-08-405-175A-9
6	51	44.3	365	4	US-09-252-991A-32327
7	51	44.3	621	4	US-09-252-991A-19125
8	49	42.6	314	4	US-09-252-991A-26728
9	48	41.7	408	4	US-09-252-991A-31571
10	47	40.9	996	4	US-09-252-991A-28596
11	47	40.9	1257	4	US-09-252-991A-17290
12	46.5	40.4	416	4	US-09-252-991A-20341
13	46	40.0	207	4	US-09-252-991A-28823
14	46	40.0	380	4	US-09-252-991A-17603
15	46	40.0	758	4	US-09-252-991A-29974
16	45	39.1	172	4	US-09-252-991A-25750
17	45	39.1	694	4	US-09-252-991A-22279
18	44	38.3	98	4	US-09-216-393B-98
19	44	38.3	123	4	US-09-489-039A-9970
20	44	38.3	287	1	US-08-146-421-5
21	44	38.3	422	4	US-09-489-039A-8150
22	44	38.3	488	4	US-09-252-991A-28535
23	44	38.3	502	4	US-09-252-991A-31240
24	44	38.3	538	4	US-09-252-991A-21622
25	44	38.3	689	4	US-09-252-991A-32669
26	44	38.3	731	4	US-09-252-991A-17180
27	44	38.3	1124	4	US-09-252-991A-26810

28	43	37.4	202	4	US-09-252-991A-25553	Sequence 25553, A
29	43	37.4	275	4	US-09-252-991A-29227	Sequence 29227, A
30	43	37.4	284	4	US-09-252-991A-32780	Sequence 32780, A
31	43	37.4	327	2	US-08-739-485-10	Sequence 10, Appl
32	43	37.4	391	3	US-09-076-359-2	Sequence 2, Appli
33	43	37.4	411	4	US-09-543-681A-7017	Sequence 7017, Ap
34	43	37.4	533	4	US-09-252-991A-23560	Sequence 23560, A
35	43	37.4	1315	4	US-09-252-991A-22746	Sequence 22746, A
36	42.5	37.0	492	4	US-09-252-991A-28339	Sequence 28339, A
37	42.5	37.0	673	4	US-09-252-991A-26458	Sequence 26458, A
38	42	36.5	208	4	US-09-252-991A-27430	Sequence 27430, A
39	42	36.5	218	4	US-09-252-991A-17692	Sequence 17692, A
40	42	36.5	272	4	US-09-252-991A-20279	Sequence 20279, A
41	42	36.5	304	4	US-09-252-991A-17069	Sequence 17069, A
42	42	36.5	331	4	US-09-252-991A-25035	Sequence 25035, A
43	42	36.5	384	4	US-09-252-991A-26093	Sequence 26093, A
44	42	36.5	536	4	US-09-252-991A-16754	Sequence 16754, A
45	42	36.5	619	4	US-09-252-991A-21585	Sequence 21585, A

#### ALIGNMENTS

RESULT 1  
US-08-405-175A-8  
; Sequence 8, Application US/08405175A  
; Patent No. 5885772  
; GENERAL INFORMATION:  
; APPLICANT: Aderem, Alan A.  
; APPLICANT: Chen, Jianmin  
; APPLICANT: Chang, Sandy  
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/405,175A  
FILING DATE: 16-MAR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-121A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 309 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
DESCRIPTION: predicted primary structure of rat MARCKS  
HYPOTHETICAL: NO  
US-08-405-175A-8

Query Match 100.0%; Score 115; DB 2; Length 309;  
Best Local Similarity 100.0%; Pred. No. 3,1e-10;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 GAQFSKTAAGKGEAAERPGEAAVA 24

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/405,175A  
FILING DATE: 16-MAR-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-121A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521



```

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

```

```

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND A
; TITLE OF INVENTION: AERUGINOSA FOR DIA
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252 9

```

; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 26728  
; LENGTH: 314  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-26728

Query Match 42.6%; Score 49; DB 4; Length 314;  
Best Local Similarity 64.3%; Pred. No. 5.9; Mismatches 2; Indels 0; Gaps 0;  
Matches 9; Conservative 3;

QY 10 KGEAAARPGGAAY 23  
|||:|||||:  
Db 30 KGQAARPGQALV 43

## RESULT 9

US-09-252-991A-31571  
; Sequence 31571, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 31571  
; LENGTH: 408  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-31571

Query Match 41.7%; Score 48; DB 4; Length 408;  
Best Local Similarity 57.9%; Pred. No. 11; Mismatches 7; Indels 0; Gaps 0;  
Matches 11; Conservative 1;

QY 6 KTAAGKGAARPGGAAY 24  
:|||||:  
Db 302 RAAGLGAARPGTAPAA 320

## RESULT 10

US-09-252-991A-28596  
; Sequence 28596, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 28596  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-28596

Query Match 40.9%; Score 47; DB 4; Length 996;  
Best Local Similarity 62.5%; Pred. No. 42; Mismatches 10; Conservative 0; Indels 6; Gaps 0;

QY 9 AKGEAAARPGGAAY 24  
|||:|||||:  
Db 696 ASGEGFAERPGRNAAA 711

## RESULT 11

US-09-252-991A-17290  
; Sequence 17290, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 17290  
; LENGTH: 1257  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-17290

Query Match 40.9%; Score 47; DB 4; Length 1257;  
Best Local Similarity 50.0%; Pred. No. 54; Mismatches 12; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGKGAARPGGAAY 24  
|:|||||:  
Db 485 GEEHLAIAAVGEEAARVGERLAA 508

## RESULT 12

US-09-252-991A-20341  
; Sequence 20341, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20341  
; LENGTH: 416  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20341

Query Match 40.4%; Score 46.5; DB 4; Length 416;  
Best Local Similarity 52.4%; Pred. No. 20; Mismatches 11; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 1 GAQFSKTAAGKGAARPGGE 20  
|||:|:  
Db 117 GAQVAQAALARGAARPEGE 137

## RESULT 13

```
US-09-252-991A-28823
; Sequence 28823, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28823
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28823

Query Match          40.0%; Score 46; DB 4; Length 207;
Best Local Similarity 45.5%; Pred. No. 11;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGFAAERPGEAA 22
   ||| : : : ||| |||
Db 80 GAQAPGRCRGATGADRPGGAA 101

RESULT 14
US-09-252-991A-17603
; Sequence 17603, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17603
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17603

Query Match          40.0%; Score 46; DB 4; Length 380;
Best Local Similarity 64.7%; Pred. No. 21;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 8 AAKGEAAERPGEAAVA 24
   ||| ||| ||| |||
Db 145 AAVGQARAERAASAAVA 161

RESULT 15
US-09-252-991A-29974
; Sequence 29974, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
```

```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29974
; LENGTH: 758
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29974

Query Match          40.0%; Score 46; DB 4; Length 758;
Best Local Similarity 56.2%; Pred. No. 45;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 9 AKGEAAERPGEAAVA 24
   ||| ||| ||| |||
Db 701 AHGESAADQGERLAA 716

Search completed: October 7, 2004, 17:36:34
Job time : 16 secs
```

**This Page Blank (uspto)**

**C;Keywords:** membrane protein



```

RESULT 7
T13161
A:Kinase anchor protein 95 - human
N:Alternate names: protein DKFZp586B1222.1
C:Species: Homo sapiens (man)
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000
C:Accession: T13161; T08786
R:Lamerdin, J.E.; McCready, P.M.; Skowronski, E.; Viswanathan, V.; Burkhardt-Schultz, K.;
J.; Danganan, L.; Erler, A.; Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.; Attix
; Quan, G.; Krommiller, B.; Arellano, A.; Saunders, C.; Ow, D.; Nolan, M.; Trong, S.
submitted to the EMBL Data Library, October 1998
A:Authors: Kobayashi, A.; Olsen, A.S.; Carrano, A.V.
A:Description: Sequence analysis of a 1.5 Mb OLFIR-rich region in 19p13.1.
A:Reference number: Z17617
A:Accession: T13161
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-692 <LAM>
A:Cross-references: EMBL:AC005785; NID:g3702290; PID:g3702291; PIDN:AAC62838.1
R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A:Reference number: Z16468
A:Accession: T08786
A:Molecule type: mRNA
A:Residues: 335-692 <KOE>
A:Cross-references: EMBL:AL050160
A:Experimental source: adult uterus; clone DKFZp586B1222
C:Genetics:
A:Map position: 19
A:Introns: 7/1; 20/1; 31/1; 124/2; 287/3; 331/1; 346/3; 358/1; 387/2; 434/3; 466/1; 509/
A:Note: DKFZp586B1222.1

Query Match 43.5%; Score 50; DB 2; Length 692;
Best Local Similarity 52.4%; Pred. No. 36;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 APAAGAGAEVVKRAAEAKQA 23
||| ||| ||| ||| |||
Db 668 APAAPAAADAEEVQTDAESKDA 688

RESULT 8
S21023
photosystem II protein psbt - upland cotton
N:Alternate names: PSII-T protein
C:Species: Gossypium hirsutum (upland cotton)
C>Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 21-Jul-2000
C:Accession: A57500; S21023
R:Kapazoglou, A.; Sagliocco, F.; Dure III, L.
J. Biol. Chem. 270, 12197-12202, 1995
A:Title: PSII-T, a new nuclear encoded luminal protein from photosystem II. Targeting ar
A:Reference number: A57500; MUID:95263572; PMID:7744870
A:Accession: A57500
A:Molecule type: DNA
A:Residues: 1-105 <KAP>
A:Cross-references: GB:X54092; NID:g18511; PIDN:CAA38027.1; PID:g18512
A:Note: submitted to the EMBL Data Library, July 1990
C:Keywords: transmembrane protein

Query Match 42.6%; Score 49; DB 2; Length 105;
Best Local Similarity 68.8%; Pred. No. 9.4;
Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 9 AGAEVVKRAAEAKQAF 24
||| ||| ||| ||| |||
Db 75 ATAEPKRGSAEAKKAY 90

RESULT 9
T04347
teosinte branched1 protein - maize (fragment)
C:Species: Zea mays (maize)

```

```

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: T04347
R:Doebley, J.; Stec, A.O.; Hubbard, L.
Nature 386, 485-488, 1997
A:Title: The evolution of apical dominance in maize.
A:Reference number: Z15300; MUID:97242406; PMID:9087405
A:Accession: T04347
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-382 <DOE>
A:Cross-references: EMBL:U94494; NID:g2051378; PIDN:AAB53060.1; PID:g2051979
A:Experimental source: strain Pioneer Inbred AP9
C:Genetics:
A:Gene: tb1
A:Map position: 1.09

Query Match 42.6%; Score 49; DB 2; Length 382;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 APAAGAGAEVVKRAAEAKQ 22
||| ||| ||| ||| |||
Db 104 ASGAEKSGASLDRAAAARK 123

RESULT 10
E83112
30S ribosomal protein S7 PA4267 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 12-Jun-2003
C:Accession: E83112
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Bri
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathog
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: E83112
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-156 <STO>
A:Cross-references: GB:AE004842; GB:AE004091; NID:g9950481; PIDN:AAG07655.1; GSPDB:GN001
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: rpsG; PA4267
C:Superfamily: ribosomal protein S7

Query Match 42.2%; Score 48.5; DB 2; Length 156;
Best Local Similarity 52.0%; Pred. No. 16;
Matches 13; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY 5 AAGGAGAEVVKRAS-----AEAKQAF 24
||| ||| ||| ||| |||
Db 127 AAEKGAAGVKKREDVHRMAEANKAF 151

RESULT 11
S26826
histone H1 - maize
C:Species: Zea mays (maize)
C>Date: 12-Nov-1997 #sequence_revision 05-Dec-1997 #text_change 23-Jul-1999
C:Accession: S26826
R:Parafimahatratra, P.; Chaubet, N.; Philipps, G.; Gigot, C.
Nucleic Acids Res. 19, 1491-1496, 1991
A:Title: Nucleotide sequence and expression of a maize H1 histone cDNA.
A:Reference number: S26826; MUID:91227140; PMID:1709276
A:Accession: S26826
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-246 <RAZ>
A:Cross-references: EMBL:X57077; NID:g22320; PIDN:CAA40362.1; PID:g22321
A:Note: the authors did not translate the codon for residue 1
C:Superfamily: histone H1

```

C:Keywords: chromosomal protein; DNA binding; nucleosome

Query Match 41.7%; Score 48; DB 2; Length 246;  
Best Local Similarity 47.6%; Pred. No. 27;  
Matches 10: Conservative 5; Mismatches 6: Indels

QY 3 APAAEGAGAEVKRASAEAKQA 23  
 ||||| : : : : : : : :  
 pb 26 APAADANAAKAKKATAPKKRA 46

## RESULT 12

cgcr-4 protein - Chlamydomonas reinhardtii (fragment)  
S19113  
C:Species: Chlamydomonas reinhardtii  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 21-Jul-2000  
C:Accession: S19113; S14466  
E:Wakarchuk, W.W.; Mueller, F.W.; Beck, C.F.  
Plant Mol. Biol. 18, 143-146, 1992  
A:Title: Two GC-rich DNA elements of Chlamydomonas reinhardtii with complex arrangements  
A:Reference number: S19113; MUID:92119224; PMID:1731966  
A:Accession: S19113  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-265 <WAK>  
A:Cross-references: EMBL:X17208; NID:g18136; PIDN:CAA35080.1; PID:g18137  
C:Genetics:  
A:Gene: cgcr-4

Query Match 41.7%; Score 48; DB 2; Length 265;  
Best Local Similarity 63.2%; Pred. No. 29;  
Matches 12: Conservative 1; Mismatches 6; Indels

QY 3 APAAEGAGAEVKKRAAEAK 21  
| | | | | : : :  
Db 183 AAAEFAAAAAKARAAAEAK 201

RESULT 13

S04157  
outer membrane protein H.8 precursor - Neisseria gonorrhoeae  
C:Species: Neisseria gonorrhoeae  
C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 08-Oct-1999  
C:Accession: S04157  
R:Baehr, W.; Gotschlich, E.C.; Hitchcock, P.J.  
Mol. Microbiol. 3, 49-55, 1989  
A:Title: The virulence-associated gonococcal H.8 gene encodes 14 tandemly repeated pentapeptides  
A:Reference number: S04157; PMID:89237900; PMID:2497299  
A:Accession: S04157  
A>Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-93 <BAE>  
A:Cross-references: GB:X14739; NID:g288430; PIDN:CRA32869.1; PID:g288431  
C:Keywords: membrane protein

Query Match 40.9%; Score 47; DB 2; Length 93;  
Best Local Similarity 52.4%; Pred. No. 16;  
Matches 11: Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAEAKQA 23  
||||| - - - - -  
Dh 56 APAAEFATATAPAAEAAATEA 76

## RESULT 14

hypothetical protein NWB1523 [imported] - Neisseria meningitidis (strain MC58 serogroup H)1072  
 Species: Neisseria meningitidis  
 C.Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
 C.Accession: H1072  
 R.Tetelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, P.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Xie, H.; Chen, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:32:03 ; Search time 8 Seconds  
(without alignments)  
156,210 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAAAGAGAEVKRAAEAKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	52	45.2	1076	YE38_SCHPO	O13992 schizosacch
2	51	44.3	88	H82_NRIGO	P11910 neisseria g
3	50	43.5	113	SS12_STRRO	P29607 streptomyce
4	50	43.5	468	HPCC_ECOLI	P42269 escherichia
5	50	43.5	692	AKA8_HUMAN	O43823 homo sapien
6	49	42.6	105	PS72_GOSHI	P31336 gossypium h
7	49	42.6	518	TPM4_DROME	P49455 drosophila
8	48.5	42.2	156	RS7_PSEAE	Q9HWD1 pseudomonas
9	48.5	42.2	156	RS7_PSEPK	Q88QN9 pseudomonas
10	48.5	42.2	156	RS7_PSESM	Q885X5 pseudomonas
11	48	41.7	245	H1_MAIZE	P23444 zea mays (m
12	47.5	41.3	155	RL5_CORGL	Q8NT20 corynebacte
13	47	40.9	110	RLA3_SCHPO	P17477 schizosacch
14	47	40.9	378	CC37_HUMAN	Q16543 homo sapien
15	47	40.9	617	41_BOVIN	Q9N179 bos taurus
16	46	40.0	110	RLA2_CRYST	O61463 cryptochic
17	46	40.0	136	RL28_HUMAN	P46779 homo sapien
18	46	40.0	136	RL28_MOUSE	P41105 mus musculu
19	46	40.0	211	NUSB_GLOVI	Q7HNZ0 gloeobacter
20	46	40.0	308	RS2_STRAW	P09788 pseudomyce
21	46	40.0	520	DH4C_PSEPU	Q82JX7 streptomyce
22	46	40.0	637	CL16_RABIT	Q9N2G5 oryctolagus
23	45.5	39.6	110	RLA4_SCHPO	P17478 schizosacch
24	45.5	39.6	111	RLA2_ARTSA	P02399 artemia sal
25	45	39.1	191	YO76_VIBCH	Q9KPY7 vibrio chol
26	45	39.1	215	ORN_MYCTU	O06174 mycobacteri
27	45	39.1	220	VG32_BPMLS	Q05241 mycobacteri
28	45	39.1	310	RS2_STRRO	Q31212 streptomyce
29	45	39.1	2567	M16B_HUMAN	Q81UG5 homo sapien
30	44.5	38.7	155	RS7_MYCLE	P30764 mycobacteri
31	44.5	38.7	155	RS7_MYCSM	P41193 mycobacteri
32	44.5	38.7	155	RS7_MYCTU	P41194 mycobacteri
33	44.5	38.7	156	RS7_MYCBO	Q53539 mycobacteri

34 44.5 38.7 156 1 RS7\_THICU  
35 44 38.3 109 1 RLAI\_TRYCR  
36 44 38.3 114 1 RLAI\_RAT  
37 44 38.3 343 1 DHAS\_CAMJE  
38 44 38.3 372 1 TOLA\_HABIN  
39 44 38.3 430 1 CIWC\_RAT  
40 44 38.3 456 1 GUNA\_MICBI  
41 44 38.3 467 1 HHCM\_HUMAN  
42 44 38.3 469 1 LEUZ\_PHOLL  
43 44 38.3 507 1 LATI\_HUMAN  
44 43.5 37.8 452 1 NIFB\_RHOCA  
45 43 37.4 107 1 SSI3\_STRCO

#### ALIGNMENTS

RESULT 1  
YE38\_SCHPO  
ID YE38\_SCHPO STANDARD; PRT; 1076 AA.  
AC O13992; O36013;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical protein C20G4.08 in chromosome I.  
GN SPAC20G4.08 OR SPAC4F10.01.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,  
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volkart G., Aert R., Robben J., Grynoprez B.,  
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
RA Shpakovski G., Ussery D., Barrell B.G., Nurse P.;  
RT "The genome sequence of Schizosaccharomyces pombe."  
CC Nature 415:871-880(2002).

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; Z98600; CAB11257.1; -;  
DR EMBL; Z98980; CAB11705.1; -;  
DR PIR; T38122; T38122.

```

DR GENE SPOMBE: SPAC20G4.08; --
KW Hypothetical protein. 119230 MW; E8FDDC3630951430 CRC64;
SQ SEQUENCE 1076 AA; 119230 MW; 88 AA.

Query Match 45.2%; Score 52; DB 1; Length 1076;
Best Local Similarity 39.1%; Pred. No. 18;
Matches 9; Conservative 7; Mismatches 0; Gaps 0;

Qy 1 GTAPAAEGAGAEVKGASAEAKQA 23
Db 95 GAKPSGTASGADVXRSDSESTE 117

RESULT 2
H82_NEIGO
ID H82_NEIGO STANDARD; PRT; 88 AA.
AC P11910;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Outer membrane protein H.8 precursor.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=485;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 700825 / FA 1090;
RX MEDLINE=89237899; PubMed=2497298;
RA Woods J.P., Spinola S.M., Strobel S.M., Cannon J.G.;
RT "Conserved lipoprotein H.8 of pathogenic Neisseria consists entirely of pentapeptide repeats."
RL Mol. Microbiol. 3:43-48(1989).
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid anchor.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X12627; CAA311145.1; --
DR PIR; S02720; S02720.
DR InterPro; IPR000437; Prok_lipoprot_S.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Repeat; Outer membrane; Lipoprotein; Signal; Palmitate.
FT SIGNAL 1 17
FT CHAIN 18 88 OUTER MEMBRANE PROTEIN H.8.
FT LIPID 18 18 N-palmitoyl cysteine.
FT LIPID 18 18 S-diacylglycerol cysteine.
FT DOMAIN 23 87 13 X 5 AA TANDEN REPEATS OF A-A-E-A-P.
SQ SEQUENCE 88 AA; 8023 MW; 834DDFAE049FCC21 CRC64;

Query Match 44.3%; Score 51; DB 1; Length 88;
Best Local Similarity 39.1%; Pred. No. 25;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 3 APAEAGAGAEVKGASAEAKQA 23
Db 46 APAEAGAGAEVKGASAEAKQA 66

RESULT 3
SSI2_STRO
ID SSI2_STRO STANDARD; PRT; 113 AA.
AC P29607; Q9R2G7;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Subtilisin inhibitor-like protein-2 (SIL-2) (SIL2).
```

```

OS Streptomyces rochei (Streptomyces parvullus).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1928;
RN [1]
RP SEQUENCE.
RC STRAIN=KCC 2283;
RX MEDLINE=94192683; PubMed=8143745;
RA Taguchi S., Kojima S., Terabe M., Miura K.-I., Momose H.;
RT "Comparative studies on the primary structures and inhibitory properties of subtilisin-trypsin inhibitors from Streptomyces."
RL Eur. J. Biochem. 220:911-918(1994).
RN [2]
RP SEQUENCE OF 1-39.
RX MEDLINE=93146392; PubMed=1490613;
RA Taguchi S., Kojima S., Kumagai I., Ogawara H., Miura K.-I., Momose H.;
RT "Isolation and partial characterization of SSI-like protease inhibitors from Streptomyces."
RL FEMS Microbiol. Lett. 78:293-297(1992).
RN [3]
RP SEQUENCE OF 1-39.
RC STRAIN=KCC 2283;
RX MEDLINE=93222542; PubMed=7763545;
RA Taguchi S., Kikuchi H., Kojima S., Kumagai I., Nakase T., Miura K.-I., Momose H.;
RT "High frequency of SSI-like protease inhibitors among Streptomyces."
RL Biosci. Biotechnol. Biochem. 57:522-524(1993).
CC -!- FUNCTION: INHIBITOR OF SUBTILISIN BPN' AND TRYPSIN.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the serine protease inhibitors SSI family.
DR PIR; S42571; S42571.
DR HSP; P01006; 3SSI.
DR InterPro; IPR000691; Strep_subst_inhib.
DR Pfam; PF00720; SSI; 1.
DR PRINTS; PR00294; SSBLININHTR.
DR ProDom; PD004028; Strep_subst_inhib; 1.
DR PROSITE; PS00999; SSI; 1.
KW Serine protease inhibitor.
FT DISULFID 35 50 BY SIMILARITY.
FT DISULFID 71 101 BY SIMILARITY.
FT ACT_SITE 73 74 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 113 AA; 11568 MW; B85B6D10787D40FC CRC64;

Query Match 43.5%; Score 50; DB 1; Length 113;
Best Local Similarity 52.6%; Pred. No. 42;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GTAPAAEGAGAEVKGASAE 19
Db 41 GTHPAADAACAEELRRAGGD 59

RESULT 4
HPCC_ECOLI
ID HPCC_ECOLI STANDARD; PRT; 468 AA.
AC P42269;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 5-carboxymethyl-2-hydroxymuconate semialdehyde dehydrogenase (EC 1.2.1.-) (CHMS dehydrogenase).
GN HPCC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C;
RX MEDLINE=95255666; PubMed=7737515;
RA Roper D.I., Stringfellow J.M., Cooper R.A.;
RT "Sequence of the hpcC and hpcG genes of the meta-fission
```



```
CC with the photosystem II complex.
CC -!- PTM: The maturation of the PSII-T precursor to its final form
CC occurs through a two step process. First, a stromal intermediate
CC is formed, which, upon translocation into the thylakoid membrane,
CC is processed to the mature protein.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X54092; CAA38027.1; -.
CC PIR; A57500; S21023.
CC Photosynthesis; Photosystem II; Chloroplast; Transit peptide;
CC Thylakoid; Membrane.
CC TRANSIT 1 77 CHLOROPLAST.
CC CHAIN 78 105 PHOTOSYSTEM II 5 kDa PROTEIN.
CC DOMAIN 62 67 POLY-ALA.
CC SEQUENCE 105 AA; 11009 MW; D0C9DA37541213DE CRC64;
CC -----
CC Query Match 42.6%; Score 49; DB 1; Length 105;
CC Best Local Similarity 68.8%; Pred. No. 5.4;
CC Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
CC -----
CC QY 9 AGAEVKRASAEAKQAF 24
CC ||| ||| ||| ||| |||
CC DB 75 ATAEPKRGSAEAKKAY 90
CC -----
CC RESULT 7
CC ID TPW4 DROME STANDARD; PRT; 518 AA.
CC AC P49455; P49456; Q24425; Q24426;
CC DT 01-FEB-1996 (Rel. 33, Created)
CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Tropomyosin I, isoforms 33/34 (Tropomyosin II).
CC TM1 OR TM11.
CC OS Drosophila melanogaster (Fruit fly).
CC OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
CC OC Ephydroidea; Drosophilidae; Drosophila.
CC OC NCBI_TaxID=7227;
CC RN [1]
CC RP SEQUENCE FROM N.A. (ISOFORMS 33 AND 34).
CC RC TISSUE=Embryo, and Pupae;
CC RX MEDLINE=89127197; PubMed=2851721;
CC RA Hanke P.D., Storti R.V.;
CC RT "The Drosophila melanogaster tropomyosin II gene produces multiple
CC RT proteins by use of alternative tissue-specific promoters and
CC RT alternative splicing."
CC RL Mol. Cell. Biol. 8:3591-3602(1988).
CC RN [2]
CC RP SEQUENCE FROM N.A. (ISOFORMS 33 AND 34).
CC RC STRAIN=Oregon-R; TISSUE=Pupae;
CC RX MEDLINE=87064486; PubMed=3097506;
CC RA Karlik C.C., Fyrberg E.A.;
CC RT "Two Drosophila melanogaster tropomyosin genes: structural and
CC RT functional aspects."
CC RL Mol. Cell. Biol. 6:1965-1973(1986).
CC CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=5;
CC Comment=Additional isoforms seem to exist;
CC Name=33; Synonyms=9C;
CC IsoID=P49455-1; Sequence=Displayed;
CC Name=Muscle; Synonyms=9D;
CC IsoID=P06754-1; Sequence=External;
CC Name=Non-muscle; Synonyms=Cytoskeletal;
CC IsoID=P06754-2; Sequence=External;
CC Name=9A;
```

```
CC IsoID=P06754-3; Sequence=External;
CC Name=34; Synonyms=9B;
CC IsoID=P49455-2; Sequence=VSP 006623, VSP 006624, VSP 006625;
CC -!- TISSUE SPECIFICITY: Both isoforms are only expressed in indirect
CC flight muscles.
CC -!- DEVELOPMENTAL STAGE: Both isoforms are expressed during pupal and
CC adult stages.
CC -!- DOMAIN: The molecule is in a coiled coil structure. The sequence
CC exhibits a prominent seven-residues periodicity.
CC -!- SIMILARITY: Belongs to the tropomyosin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X76208; CAA53800.1; -.
CC EMBL; X76208; CAA53801.1; -.
CC EMBL; K02620; AAA28967.1; ALT_SEQ.
CC EMBL; L00355; AAA28967.1; JOINED.
CC EMBL; L00356; AAA28967.1; JOINED.
CC EMBL; L00357; AAA28967.1; JOINED.
CC EMBL; L00358; AAA28967.1; JOINED.
CC EMBL; L00359; AAA28967.1; JOINED.
CC EMBL; L00360; AAA28967.1; JOINED.
CC EMBL; L00362; AAA28967.1; JOINED.
CC EMBL; L00362; AAA28967.1; JOINED.
CC EMBL; M12840; AAA28967.1; JOINED.
CC EMBL; K02621; AAA28968.1; -.
CC EMBL; M12840; AAA28968.1; JOINED.
CC EMBL; L00355; AAA28968.1; JOINED.
CC EMBL; L00356; AAA28968.1; JOINED.
CC EMBL; L00357; AAA28968.1; JOINED.
CC EMBL; L00358; AAA28968.1; JOINED.
CC EMBL; L00359; AAA28968.1; JOINED.
CC EMBL; L00360; AAA28968.1; JOINED.
CC EMBL; L00362; AAA28968.1; JOINED.
CC FlyBase: FBgn003721; Tm1.
CC GO; GO:0045451; P:pole plasm oskar mRNA localization; IMP.
CC InterPro; IPR000533; Tropomyosin.
CC Pfam; PF00261; Tropomyosin; 1.
CC PRINTS; PR00194; TROPOMYOSIN.
CC PROSITE; PS00326; TROPOMYOSIN; 1.
CC Muscle protein; Cytoskeleton; Actin-binding; Coiled coil;
CC Alternative splicing; Multigene family;
CC COILED COIL (POTENTIAL).
CC DOMAIN 14 267
CC DOMAIN 287 518
CC VARSPLIC 259 293
CC -----
CC VARSPLIC 300 367
CC -----
CC VARSPLIC 391 518
CC -----
CC CONFLICT 106 114
CC CONFLICT 119 119
CC CONFLICT 183 183
CC CONFLICT 199 199
CC CONFLICT 503 503
CC SEQUENCE 518 AA; 54558 MW; 153D0872CF9DB6EA CRC64;
```

Query Match 42.6%; Score 49; DB 1; Length 518;  
Best Local Similarity 57.1%; Pred. No. 23;  
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAAGAGAEVGRASAEAKQA 23  
DB 476 APAAGAAFAEAPAEATRAEA 496

## RESULT 8

RS7\_PSEAF STANDARD; PRT; 156 AA.  
ID RS7\_PSEAF  
AC Q9HWD1;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 30S ribosomal protein S7.  
GN RPSG OR PA4267.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,  
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an  
RT opportunistic pathogen."  
RL Nature 406:959-964 (2000).

CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds  
CC directly to 16S rRNA where it nucleates assembly of the head  
CC domain of the 30S subunit. Is located at the subunit interface  
CC close to the decoding center, probably blocks exit of the E-site  
CC tRNA (By similarity).  
CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts proteins S9  
CC and S11 (By similarity).  
CC -!- SIMILARITY: Belongs to the S7P family of ribosomal proteins.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

CC EMBL; AE004842; AAC07655.1; -;  
CC F1R; E83112; E83112.  
CC HSP; P22744; LHUS.

DR HAMAP; MF 00480; -; 1.  
DR InterPro; IPR00235; Ribosomal\_S7.  
DR InterPro; IPR005717; Ribosomal\_S7\_b/o.  
DR Pfam; PF00177; Ribosomal\_S7; 1.  
DR Pfam; PD000817; Ribosomal\_S7; 1.  
DR TIGRFAMs; TIGR01029; rpsg\_bact; 1.  
DR PROSITE; PS00052; rpsg\_bact; 1.  
DR Ribosomal protein; RNA-binding; tRNA-binding;  
KW Complete proteome.  
SQ SEQUENCE 156 AA; 17504 MW; B747FFDB710DDC14 CRC64;

Query Match 42.2%; Score 48.5; DB 1; Length 156;  
Best Local Similarity 52.0%; Pred. No. 9;  
Matches 13; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY 5 AAEAGAGAEVGRASAEAKQA 24  
DB 127 AAEAGAAVKKREDVHRMAEANKAF 151

## RESULT 9

RS7\_PSEPK STANDARD; PRT; 156 AA.  
ID RS7\_PSEPK  
AC Q88QN9;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE 30S ribosomal protein S7.  
GN RPSG OR PP0450.  
OS Pseudomonas putida (strain KT2440).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
OC Pseudomonadaceae; Pseudomonas.  
OX NCBI\_TaxID=160488;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22423060; PubMed=12534463;  
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,  
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,  
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,  
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,  
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,  
RA Utterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,  
RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,  
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,  
RA Fraser C.M.,  
RT "Complete genome sequence and comparative analysis of the  
RT metabolically versatile Pseudomonas putida KT2440."  
RL Environ. Microbiol. 4:799-808 (2002).  
CC -!- FUNCTION: One of the primary rRNA binding proteins, it binds  
CC directly to 16S rRNA where it nucleates assembly of the head  
CC domain of the 30S subunit. Is located at the subunit interface  
CC close to the decoding center, probably blocks exit of the E-site  
CC tRNA (By similarity).  
CC -!- SUBUNIT: Part of the 30S ribosomal subunit. Contacts proteins S9  
CC and S11 (By similarity).  
CC -!- SIMILARITY: Belongs to the S7P family of ribosomal proteins.  
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

DR EMBL; AE016775; AAN66080.1; -;  
DR TIGR; PP0450; -;  
DR HAMAP; MF 00480; -; 1.  
DR InterPro; IPR000235; Ribosomal\_S7.  
DR Pfam; PF00177; Ribosomal\_S7; 1.  
DR PROSITE; PS00052; RIBOSOMAL\_S7; 1.  
DR Ribosomal protein; RNA-binding; tRNA-binding;  
KW Complete proteome.  
SQ SEQUENCE 156 AA; 17578 MW; 22B47BC15B4B921A CRC64;

Query Match 42.2%; Score 48.5; DB 1; Length 156;  
Best Local Similarity 52.0%; Pred. No. 9;  
Matches 13; Conservative 2; Mismatches 5; Indels 5; Gaps 1;

QY 5 AAEAGAGAEVGRASAEAKQA 24  
DB 127 AAEAGAAVKKREDVHRMAEANKAF 151

## RESULT 10

RS7\_PSESM STANDARD; PRT; 156 AA.  
ID RS7\_PSESM  
AC Q88X5;  
DT 15-MAR-2004 (Rel. 43, Created)  
DT 15-MAR-2004 (Rel. 43, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)



```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP005275; BAB97887.1; -.
DR HAMAP; MF 00480; -.
DR InterPro; IPR000235; Ribosomal_S7.
DR InterPro; IPR005717; Ribosomal_S7_b/o.
DR Pfam; PF00177; Ribosomal_S7; 1.
DR ProDom; PD000817; Ribosomal_S7; 1.
DR TIGRFAMs; TIGR01029; rpsG bact; 1.
DR PROSITE; PS00052; RIBOSOMAL_S7; 1.
KW Ribosomal protein; RNA-binding; tRNA-binding;
KW Complete proteome;
KW SEQUENCE 155 AA; 17481 MW; C5AD347939343B70 CRC64;
SQ
Query Match 41.3%; Score 47.5; DB 1; Length 155;
Best Local Similarity 52.0%; Pred. No. 12;
Matches 13; Conservative 1; Mismatches 6; Indels 5; Gaps 1;
QY 5 AREGAGAEVKRAS-----AEAKQAF 24
DB 126 AANGLGASVKRRRDTKMAEANRAF 150
|| | | | | | | | | | | | | |
|| | | | | | | | | | | | | |
RESULT 13
RLA3 SCHPO
ID RLA3 SCHPO STANDARD; PRT; 110 AA.
AC P17477;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 60S acidic ribosomal protein P1-alpha 3 (A3).
GN RPA3 OR SPBC3B9.13C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90220620; PubMed=2325655;
RA Beltrame M., Bianchi M.E.;
RT "A gene family for acidic ribosomal proteins in Schizosaccharomyces
RL pombe: two essential and two nonessential genes.";
RL Mol. Cell. Biol. 10:2341-2348(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Guillermo R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagers K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymprez B.,
RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,

```

```

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Certutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Plays an important role in the elongation step of
CC protein synthesis.
CC -!- SUBUNIT: P1 and P2 exist as dimers at the large ribosomal subunit.
CC -!- MISCELLANEOUS: Yeasts contain 4 individual small ribosomal A
CC proteins (RPA) which can be classified into two couples of similar
CC but not identical sequences. Each couple is distinctly related to
CC one of the two A proteins present in multicellular organisms.
CC -!- MISCELLANEOUS: Rpa3 and rpa4 are essential for cell survival,
CC whereas rpa1 and rpa2 are not.
CC -!- SIMILARITY: Belongs to the L12P family of ribosomal proteins.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M331139; AAA35336.1; -.
DR EMBL; AL022070; CAA17793.1; -.
DR PIR; C34715; R6BYP3.
DR GenDB SPombe; SPBC3B9.13c; -.
DR InterPro; IPR001813; Ribosomal_60S.
DR Pfam; PF00428; 60s ribosomal; 1.
KW Ribosomal protein; Phosphorylation; Multigene family.
SQ SEQUENCE 110 AA; 11171 MW; 1734AC9779F5A891 CRC64;
QY 3 APAAGGAGAEVKRASAEA 20
DB 68 APAAGGAGAPAAAAGEA 85
|| | | | | | | | | | | | | |
|| | | | | | | | | | | | | |
Query Match 40.9%; Score 47; DB 1; Length 110;
Best Local Similarity 61.1%; Pred. No. 10;
Matches 11; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 3 APAAGGAGAEVKRASAEA 20
DB 68 APAAGGAGAPAAAAGEA 85
|| | | | | | | | | | | | | |
|| | | | | | | | | | | | | |
RESULT 14
CC37 HUMAN
ID CC37_HUMAN STANDARD; PRT; 378 AA.
AC Q16543;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hsp90 co-chaperone Cdc37 (Hsp90 chaperone protein kinase-targeting
DE subunit) (p50Cdc37).
GN CDC37.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=96258250; PubMed=8666233;
RA Stepanova L., Leng X., Parker S.B., Harper J.W.;
RT "Mammalian p50Cdc37 is a protein kinase-targeting subunit of Hsp90
RT that binds and stabilizes Cdk4.";
RL Genes Dev. 10:1491-1502(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX Dai K., Kobayashi R., Beach D.;
RT "Physical interaction of mammalian CDC37 with CDK4.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Lymph, and Placenta;
RX MEDLINE=22388257; PubMed=12477932;

```







FT DOMAIN 467 617 CARBOXYL-TERMINAL (CTD).  
SQ SEQUENCE 617 AA; 69257 MW; 5B49D5008AD900FB CRC64;  
Query Match 40.9%; Score 47; DB 1; Length 617;  
Best Local Similarity 54.5%; Pred. No. 51;  
Matches 12; Conservative 2; Mismatches 6; Indels 2; Gaps 1;  
QY 3 APAAGA--GAEVKRAAEAKQ 22  
Db 358 SPAEGGVGAPVKAKQETVQ 379

Search completed: October 7, 2004, 17:34:18  
Job time : 10 secs

**This Page Blank (uspto)**

RESULT 2	
Q7XEN1	
ID Q7XEN1	PRELIMINARY;
AC Q7XEN1;	PRT; 406 AA.
DT 01-OCT-2003	(TReMBLrel. 25, Created)
DT 01-OCT-2003	(TReMBLrel. 25, Last sequence up

```
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative proline transporter 1.
GN OSUNBA003i20.1.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA "The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
RL chromosome 10."
RL Science 300:1566-1569(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buehl C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017094; AAP33757.1; E640F1701867EE72 CRC64;
SQ SEQUENCE 406 AA; 44212 MW; E640F1701867EE72 CRC64;

Query Match 46.1%; Score 53; DB 10; Length 406;
Best Local Similarity 52.6%; Pred. No. 43;
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVVKRASAEE 19
   |||||:|:|:|:|:|:|
DB 6 GGTPEEGGAGADVEKAAAE 24

RESULT 3
Q9ATU0 PRELIMINARY; PRT; 349 AA.
ID Q9ATU0
AC Q9ATU0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Teosinte branched1 protein (Fragment).
GN TBI.
OS Zea diploerennis (Diploperennial teosinte).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4576;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21165336; PubMed=11264415;
RA Lukens L., Doebley J.;
RT "Molecular evolution of the teosinte branched gene among maize and
RT related grasses."
RL Mol. Biol. Evol. 18:627-638(2001).
DR EMBL; AF322117; AAK37479.1; -.
DR InterPro; IPR005333; TCP.
DR Pfam; PF03634; TCP; 1.
FT NON_TER 1
FT NON_TER 349
SQ SEQUENCE 349 AA; 36858 MW; F72D22E76A1999B4 CRC64;

Query Match 45.2%; Score 52; DB 10; Length 349;
Best Local Similarity 55.0%; Pred. No. 50;
Matches 11; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 APAAGAGAEVVKRASAEEQ 22
   |||||:|:|:|:|:|:|
DB 99 ASGAGAGASLDRAAAARK 118

RESULT 4
Q69371 PRELIMINARY; PRT; 540 AA.
ID Q69371
AC Q69371;
```

```
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glycoprotein GE.
OS Cercopithecine herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae.
OX NCBI_TaxID=10317;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93298054; PubMed=8390827;
RA Eberle R., Zhang M., Black D.;
RT "Gene mapping and sequence analysis of the unique short region of the
RT simian herpesvirus SA 8 genome."
RL Arch. Virol. 130:391-411(1993).
DR EMBL; AF449714; AAA46180.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR003404; Herpes_glycopE.
DR InterPro; IPR007110; IG-like.
DR Pfam; PF02480; Herpes_GE; 1.
SQ SEQUENCE 540 AA; 58049 MW; BC34E60B2F3392EE CRC64;

Query Match 45.2%; Score 52; DB 12; Length 540;
Best Local Similarity 52.2%; Pred. No. 78;
Matches 12; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVVKRASAEEAKQA 23
   |||||:|:|:|:|:|:|
DB 172 GPPPLAEGAGAEAGAAATRAPAA 194

RESULT 5
Q8S5P3 PRELIMINARY; PRT; 183 AA.
ID Q8S5P3
AC Q8S5P3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSJNB0059B20.23 OR OSJNAA001109.18.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA McCombie W.R., de la Bastide M., Spiegel L., Preston R., Kirchoff K.,
RA Kuit K., Nascimento L., Zutavern T., Balija V., Bell M., Baker J.,
RA Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,
RA O'Shaughnessy A., Palmer L., Dedhia N.;
RT "Genomic sequence for Oryza sativa, Nipponbare strain, clone
RT OSJNB0059B20, from chromosome 10, complete sequence."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Yang T.J., Nah G., Soderlund C., Chen M., Kim H.-R.,
RA Rambo T., Saski C., Henry D., Oates R., Simmons J.;
RT "Rice Genomic Sequence."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
RT chromosome 10."
RL Science 300:1566-1569(2003).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buehl C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
```

```

DR EMBL; AC108884; AAM01141.1; -.
DR EMBL; AC122144; AAM44895.1; -.
DR EMBL; AE017090; AAP53610.1; -.
DR Gramene; O8S5P3; -.
KW Hypothetical protein.
SQ SEQUENCE 183 AA; 19747 MW; 8E83AE03BEE44AB1 CRC64;

Query Match          44.3%; Score 51; DB 10; Length 183;
Best Local Similarity 45.5%; Pred. No. 35;
Matches 10; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVKRASAEAKQ 22
DB 38 GDRPTAEGGGAATKRCGREGER 59

RESULT 6
O8N2Q3 PRELIMINARY; PRT; 316 AA.
AC O8N2Q3;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein FL90054 (Fragment).
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Isogai T., Ota T., Nishikawa T., Hayashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuko Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RA "NEO human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074535; BAC11045.1; -.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
KW Hypothetical protein.
FT NON_TER 316 316
SQ SEQUENCE 316 AA; 34030 MW; 167FC61940485600 CRC64;

Query Match          44.3%; Score 51; DB 4; Length 316;
Best Local Similarity 60.0%; Pred. No. 62;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 TAPAAEGAGAEVKRASAEAK 21
DB 16 SAAGAGAGAPAAASAEPK 35

RESULT 7
Q9GZF7 PRELIMINARY; PRT; 460 AA.
AC Q9GZF7;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN C18H7.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;

RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RT Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RT Tin-Wollam A., Fronicke W.;
RT "The sequence of C. elegans cosmid C18H7.";
RT Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RT Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF067607; AAF98609.1; -.
DR PIR; T33110; T33110.
DR WormPep; C18H7.3; CE17428.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR002486; Col_cuticle_N.
DR Pfam; PF01391; Collagen; 2.
DR Pfam; PF01484; Col_cuticle_N; 1.
KW Hypothetical protein.
SQ SEQUENCE 460 AA; 41016 MW; BF99CC80770E2202 CRC64;

Query Match          44.3%; Score 51; DB 5; Length 460;
Best Local Similarity 57.1%; Pred. No. 91;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAEAKQA 23
DB 308 APAAEGAGGAGPAGAAAPDA 328

RESULT 8
Q8ZQ48 PRELIMINARY; PRT; 488 AA.
AC Q8ZQ48;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE 4-hydroxyphenylacetate catabolism (EC 1.2.1.60).
GN HPAE OR STM1102.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RT Nature 413:852-856 (2001).
RL EMBL; AE008747; AAL20034.1; -.
DR GO; GO:0018480; F:5-carboxymethyl-2-hydroxy-5-oxo-2-pyridinecarboxaldehyde; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyde_dehydr.
DR Pfam; PF00171; aldehyd; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 488 AA; 53039 MW; 7CEBF3AB95FBA8CF CRC64;

Query Match          44.3%; Score 51; DB 16; Length 488;
Best Local Similarity 54.5%; Pred. No. 96;

```

Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy 3 APAEAGAGAEVVKRASAEAK 24

Db 31 AEVASGGEAEVQNQAAAEAF 52

RESULT 9

Q9HOT8  
ID Q9HOT8 PRELIMINARY; PRT; 589 AA.  
AC Q9HOT8;  
DT 01-MAR-2001 (TRENBLrel. 16, Created)  
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN DKFZP564L023.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Waubutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL136643; CAB66578.1; -.  
DR HSSP; P02248; 1TBE.  
DR InterPro; IPR006636; STIL.  
DR InterPro; IPR000449; UBA domain.  
DR InterPro; IPR000626; Ubiquitin.  
DR Pfam; PF00627; UBA; 1.  
DR Pfam; PF00240; ubiquitin; 1.  
DR SMART; SM00727; STIL; 4.  
DR SMART; SM00165; UBA; 1.  
DR SMART; SM00213; UBQ; 1.  
DR PROSITE; PS50030; UBA; 1.  
DR PROSITE; PS50053; UBIQUITIN\_2; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 589 AA; 62489 MW; F41E0F394862FF7C CRC64;

Query Match 44.3%; Score 51; DB 4; Length 589;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 2 TAPAEAGAGAEVVKRASAEAK 21

Db 16 SAAGAEGAGAPAAASAEPK 35

RESULT 10

Q9UMX0  
ID Q9UMX0 PRELIMINARY; PRT; 589 AA.  
AC Q9UMX0; Q9HA25;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Ubiquilin (PILC-1).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=20530664; PubMed=11076969;  
RA Mah A.L., Perry G., Smith M.A., Monteiro M.J.;  
RT "Identification of ubiquilin, a novel presenilin interactor that  
increases presenilin protein accumulation.";  
RL J. Cell Biol. 151:847-862(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Mah A.L., Monteiro M.J.;  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
RN [3]

RP SEQUENCE FROM N.A.  
RX MEDLINE=20437281; PubMed=10983987;  
RA Kleijnen M.F., Shih A.H., Zhou P., Kumar S., Soccio R.E.,  
RA Kedersha N.L., Gill G., Howley P.M.;  
RT "The nPLIC proteins may provide a link between the ubiquitination  
machinery and the proteasome.";  
RL Mol. Cell 6:409-419(2000).  
DR EMBL; AF176069; AAD49751.3; -.  
DR EMBL; AF293384; AAG02473.1; -.  
DR HSSP; P02248; 1TBE.  
DR InterPro; IPR006636; STIL.  
DR InterPro; IPR000449; UBA domain.  
DR InterPro; IPR000626; Ubiquitin.  
DR Pfam; PF00627; UBA; 1.  
DR Pfam; PF00240; ubiquitin; 1.  
DR SMART; SM00727; STIL; 4.  
DR SMART; SM00165; UBA; 1.  
DR SMART; SM00213; UBQ; 1.  
DR PROSITE; PS50030; UBA; 1.  
DR PROSITE; PS50053; UBIQUITIN\_2; 1.  
SQ SEQUENCE 589 AA; 62519 MW; 8B4756B6113B7025 CRC64;

Query Match 44.3%; Score 51; DB 4; Length 589;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 2 TAPAEAGAGAEVVKRASAEAK 21

Db 16 SAAGAEGAGAPAAASAEPK 35

RESULT 11

Q9H3R4  
ID Q9H3R4 PRELIMINARY; PRT; 589 AA.  
AC Q9H3R4;  
DT 01-MAR-2001 (TRENBLrel. 16, Created)  
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE DA41.  
GN DA41.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RA Hanaka E.;  
RT "human DA41.";  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB035275; BAB20436.1; -.  
DR HSSP; P02248; 1TBE.  
DR InterPro; IPR006636; STIL.  
DR InterPro; IPR000449; UBA domain.  
DR InterPro; IPR000626; Ubiquitin.  
DR Pfam; PF00627; UBA; 1.  
DR Pfam; PF00240; ubiquitin; 1.  
DR SMART; SM00727; STIL; 4.  
DR SMART; SM00165; UBA; 1.  
DR SMART; SM00213; UBQ; 1.  
DR PROSITE; PS50030; UBA; 1.  
DR PROSITE; PS50053; UBIQUITIN\_2; 1.  
SQ SEQUENCE 589 AA; 62443 MW; 5E3CA56C57F61A25 CRC64;

Query Match 44.3%; Score 51; DB 4; Length 589;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

Qy 2 TAPAEAGAGAEVVKRASAEAK 21

Db 16 SAAGAEGAGAPAAASAEPK 35

```

RESULT 12
Q9KKR7
ID Q9KKR7 PRELIMINARY; PRT; 666 AA.
AC Q9KKR7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Methyl-accepting chemotaxis protein.
GN VCA1034.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E1 Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eise J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Gill S.R., Nelson K.E., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483 (2000).
DR EMBL; AF004429; AAF96929.1; -.
DR F1R; D82386; D82386.
DR HSP; P02942; LQ07.
DR TIGR; VCA1034; -.
DR GO; GO:0016020; C-membrane; IEA.
DR GO; GO:0004871; F-signal transducer activity; IEA.
DR GO; GO:0006935; P-chemotaxis; IEA.
DR GO; GO:0007165; P-signal transduction; IEA.
DR InterPro; IPR004089; Chmtaxis_trans.
DR InterPro; IPR003660; HAMP.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00283; MA; 1.
DR PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS50885; HAMP; 1.
KW Complete proteome.
SQ SEQUENCE 666 AA; 73398 MW; AE89ABDC756A9401 CRC64;

Query Match 44.3%; Score 51; DB 16; Length 666;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 APAEAGAGAEVKRASAEAKQ 22
Db 441 ARSTDGKDEVDRAAEAKQ 460

RESULT 13
Q9LON7
ID Q9LON7 PRELIMINARY; PRT; 956 AA.
AC Q9LON7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein SCO4254.
GN SCO4254 OR SCD8A.27.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3 (2);
RX Seeger K.J., Harris D.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3 (2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3 (2);
RX MEDLINE=97000351; PubMed=8843436;
RA Kedenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96 (1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3 (2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147 (2002).
DR EMBL; AL939119; CAB77348.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 956 AA; 101195 MW; F67B04C6D615F00B CRC64;

Query Match 44.3%; Score 51; DB 16; Length 956;
Best Local Similarity 52.4%; Pred. No. 1.9e+02;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 GTAPAEAGAGAEVKRASAEAK 21
Db 500 GVARAADGVGHGAKKASAKTK 520

RESULT 14
Q828Y4
ID Q828Y4 PRELIMINARY; PRT; 224 AA.
AC Q828Y4
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative peptidoglycan-binding protein.
GN SAV6527.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Onose T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis; deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;

```

RT "Complete genome sequence and comparative analysis of the industrial  
RT microorganism *Streptomyces avermitilis*,"  
RL *Nat. Biotechnol.* 21:526-531(2003).

DR EMBL: AP005047; BAC74238.1; -.

DK EMBL; AF005047; DACC4238.1;  
DR GO: GO:0016998: P:cell wall catabolism: IEA;

```
DR; GU:0016958; F:CELL wa  
InterPro: TPR002482; JvSM
```

DR InterPro; IPR002482; L  
DR Pfam: PF01476; TrEM: 1

DR PFam; PF01476; LysM; 1.

DR SMART; SM00257; LYSM; 1.

KW	Complete proteome.
SQ	SEQUENCE 224 AA; 22577 MW; 2244CB5EC7AFB37E CRC64;

Query Match 43.9%: Score 50.5: DB 16: Length 224:

Query Match	Best Local Similarity	Pred No	Score
43.3%	40.6%	51	30.57

	Best Local Similarity	40.6%	Pred. No. 51;
Matched	13	Connections	4: Mismatched
			5: Indels
			9: Gaps
			1:

1 GTAPAAEG-----AGAEVKRASAEAKOA 23

QY I GIAPAAEG-----AGAEVKKASHEANQA 23  
|:||||| ||: ::| ||||  
pb 124 GSAPAAAGTGASGGSSSAGSATTESAASAKOA 155

## RESULT 15

RESULTS  
Q8RKT0  
ID Q8RKT0  
PRELIMINARY: PRT: 158 AA.

ID Q8RKTU PRELIMINARY; FBI; 138 AA

AC Q8RKTU: Q8RKTU: FBI; 138 AA

Q8RA10;  
AC 01-JUN-2003 (T-EMP1 va) 31 (Created)

DT	01-JUN-2002	(TremBLrel. 21, Created)
DT	01-JUN-2002	(TremBLrel. 21, Created)
DT	01-JUN-2002	(TremBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)

DE Putat.

GN PKAD.

OS Frankia sp. (strain EuK1).

OS *Frankia* sp. (SCLATH EUK1).  
 QC Bacteria: Actinobacteria: Actinob

Bacteria; Actinobacteria; Frank

20 Francs

OX NCBI\_TaxID=47227;

RN [1]

RP SEQUENCE FROM N.A.

RA Oh C.J., An C.S.;

RT. Submitted (MAR-2002) to the EMBL/GenBank.

RU  
DP  
SUBMIT  
EMBT.: A

DR EMBL; AY091786;  
DB CO: 0016301;

DR GO; GO:0016301; F:kinase activity; IEA.

Query Match

Query Match	Best Local Similarity	Pred No.	Score	DE
43.38	61.18	41	30	41

Best Local Similarity	61.18;	Pred. NO. 41;
Matches	11.	Consecutive 1.
		Mismatches 6.
		Indels 0.
		Gaps 0.

1 GTAPAAEGAGAEVKRASA 18  
Ov

QY I G I A F A A E G A G A E V K A S H I S

Search completed: October 7, 2004. 17:35:35

search completed: Oct  
Job time: 38.5 sec



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:32:02 ; Search time 50.5 Seconds  
(without alignments)  
134.280 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAAEGAGAEVKTASAEAKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	24	6	Abp97766 Scrambled
2	53	46.1	487	6	Abu22475 Protein e
3	53	46.1	589	4	Aab74670 Human pro
4	51	44.3	79	6	Abp79644 N. gonorr
5	51	44.3	316	4	Aam93216 Human pol
6	51	44.3	409	6	Abu17304 Protein e
7	51	44.3	412	7	Ada28111 Human NTR
8	51	44.3	422	6	Ada33525 Acinetoba
9	51	44.3	488	6	Abu47184 Protein e
10	51	44.3	589	4	Aab95205 Human pro
11	51	44.3	589	4	Aab94311 Human pro
12	51	44.3	589	4	Abu52717 Human bra
13	51	44.3	589	4	Aab62194 Human SAP
14	51	44.3	589	4	Aam94008 Human sto
15	51	44.3	589	7	Ades9377 Human pro
16	51	44.3	604	4	Abg16557 Novel hum
17	50	43.5	113	2	Aaw75885 Peptide 1
18	50	43.5	113	3	Aay78865 Streptomy
19	50	43.5	116	3	Aab37425 Human sec
20	50	43.5	116	3	Aab37426 Human sec
21	49.5	43.0	74	5	Abp35189 Human dea
22	49.5	43.0	280	4	Aau61421 Propionib
23	49.5	43.0	280	6	Aam57940 Propionib
24	49	42.6	418	6	Ada34413 Acinetoba
25	49	42.6	518	4	Aau32416 Novel hum

ALIGNMENTS

RESULT 1

ABP97766  
ID ABP97766 standard; peptide; 24 AA.

XX  
AC ABP97766;

XX  
DT 11-AUG-2003 (first entry)

XX  
DE Scrambled myristoylated N-terminal sequence (MANS) peptide.

XX  
KW Myristoylated N-terminal sequence; MANS; MARCKS; mucus secretion;  
inflammatory mediator; inflammation; respiratory disease; asthma;  
chronic bronchitis; chronic obstructive pulmonary disease; COPD;  
bowel disease; irritable bowel syndrome; Crohn's disease;  
ulcerative colitis; skin disease; rosacea; eczema; psoriasis; acne;  
autoimmune disease; pain; arthritis; cystic fibrosis.

XX  
OS Synthetic.

XX  
FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal myristate chain attached"

XX  
FT WO2003000027-A2.

XX  
XX 03-JAN-2003.

XX  
PF 26-JUN-2002; 2002WO-US022270.

XX  
PR 26-JUN-2001; 2001US-0300933P.

XX  
PA (UINC-) UNIV NORTH CAROLINA STATE.

XX  
PI Martin LD, Adler KB, Li Y;

XX  
DR WPI; 2003-278239/27.

XX  
PT Method of regulating inflammation comprises administering a composition comprising a MANS peptide or an active fragment thereof.

XX  
PS Disclosure; Page 26; 54pp; English.

XX  
CC The present sequence represents a scrambled myristoylated N-terminal  
sequence (MANS) peptide. It is used as a control peptide, in the method  
of the invention. The MANS peptide is identical to the first 24 amino  
acids of MARCKS, and mediates insertion of MARCKS into membranes. The  
MANS peptide inhibits both mucus secretion and inflammatory mediators.  
The specification describes a method of regulating inflammation. The



PT autoimmune/inflammatory disorders such as acquired immunodeficiency  
PT syndrome, Cushing's disease, Addison's disease and cell proliferative  
PT disorders such as cancer.  
PS Claim 1; Page 93-95; 134pp; English.  
XX  
XX AAF81714 to AAF81740 encode the human proteases and protease inhibitors  
CC (PPIMs) given in AAB74668 to AAB74694. The PPIMs can have activities such  
CC as: anti-human immunodeficiency virus (HIV); antidiabetic; antithyroid;  
CC immunostimulant; immunomodulator; anti-inflammatory; immunosuppressive;  
CC nephrotropic; antigout; thyromimetic; cytostatic; antibacterial;  
CC fungicide; protozoacide; antiarteriosclerotic; antiatherosclerotic;  
CC viricide; antipsoriatic; and hepatotropic. PPIM polynucleotide and  
CC protein sequences can be used in the diagnosis, treatment and prevention  
CC of autoimmune/inflammatory disorders such as AIDS, DiGeorge's syndrome,  
CC severe combined immunodeficiency disease (SCID), Chediak-Higashi  
CC syndrome, Cushing's disease, Addison's disease, autoimmune thyroiditis,  
CC Crohn's disease, diabetes mellitus, Good pasture's syndrome, gout,  
CC Grave's diseases, Hashimoto's thyroiditis, Sjogren's syndrome, Werner's  
CC syndrome, viral, bacterial, fungal, parasitic, protozoal, and helminthic  
CC infections and cell proliferative disorder such as arteriosclerosis,  
CC atherosclerosis, cirrhosis, hepatitis, psoriasis and cancer. PPIM  
CC polynucleotide sequences can be used in somatic or germline gene therapy  
CC and in diagnosis of diseases. They can also be used in generating  
CC hybridisation probes useful in mapping the naturally occurring genomic  
CC sequences and in molecular biology techniques  
XX  
XX Sequence 589 AA;

Query Match 46.1%; Score 53; DB 4; Length 589;  
Best Local Similarity 60.0%; Pred. No. 30;  
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 TAPAEAGAGAEVVKRASAARAK 21  
Db 16 SAAGAEAGAPAAASADAK 35

RESULT 4  
ABP79644  
ID ABP79644 standard; protein; 79 AA.  
XX  
XX AC ABP79644;

DT 07-MAR-2003 (first entry)  
DE N. gonorrhoeae amino acid sequence SEQ ID 5818.  
XX Antibacterial; infection; vaccine; gene therapy.  
XX  
XX Neisseria gonorrhoeae.

WO200279243-A2.

10-OCT-2002.

12-FEB-2002; 2002WO-IB002069.

12-FEB-2001; 2001GB-00003424.

(CHTR-) CHIRON SPA.

Fontana MR, Pizza M, Masignani V, Monaci E;

WPI; 2003-058415/05.

N-PSDB; ABZ40614.

New protein from Neisseria gonorrhoeae, useful for the manufacture of a  
medicament for treating or preventing N. gonorrhoeae infection.

Disclosure; Page 612; 815pp; English.

The present invention relates to proteins from Neisseria gonorrhoeae.

CC Also disclosed are the nucleic acid molecules encoding the proteins and  
CC antibodies that specifically bind to the proteins. The composition  
CC comprising the protein, nucleic acid or antibody is useful for the  
CC manufacture of a medicament for treating or preventing N. gonorrhoeae  
CC infection, this may be in the form of a vaccine or gene therapy.  
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid  
CC molecules of the invention

XX Sequence 79 AA;

Query Match 44.3%; Score 51; DB 6; Length 79;  
Best Local Similarity 57.1%; Pred. No. 6.7;  
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAEAGAGAEVVKRASAARAKQA 23  
Db 37 APAEAGAGAEVVKRASAARAKQA 57

RESULT 5

AAM93216

ID AAM93216 standard; protein; 316 AA.

XX AAM93216;

DT 06-NOV-2001 (first entry)

DE Human polypeptide, SEQ ID NO: 2620.

KW Human; full length cDNA; cDNA synthesis; oligo-capping.

OS Homo sapiens.

PN EP1130094-A2.

PD 05-SEP-2001.

PF 07-JUL-2000; 2000EP-00114089.

PR 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.

PR 02-MAY-2000; 2000JP-00183765.

PA (HELI-) HELIX RES INST.

Ota T, Nishikawa T, Isogai T, Hayaishi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

WPI; 2001-524255/58.

N-PSDB; AAK94125.

830 Primers useful for synthesizing full length cDNA clones and their use  
in genetic manipulation.

Claim 8; SEQ ID NO 2620; 1380pp + Sequence Listing; English.

The invention relates to primers for synthesising full length cDNA  
clones. 830 cDNA molecules encoding a human protein have been isolated  
and nucleotide sequences of 5' - and 3' -ends of the cDNA molecules have  
been determined. Primers for synthesising the full length cDNA are useful  
for clarifying the function of the protein encoded by the cDNA. The full  
length clones were obtained by construction of full length enriched cDNA  
libraries that were synthesised by the oligo-capping method. The primers  
enable the production of the full length cDNA easily without any special  
methods. The present sequence is a polypeptide encoded by a full length  
human cDNA of the invention. Note: The sequence data for this patent did  
not form part of the printed specification, but was obtained in CD-ROM  
format directly from EPO

XX Sequence 316 AA;

Query Match 44.3%; Score 51; DB 4; Length 316;  
Best Local Similarity 60.0%; Pred. No. 30;

Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
QY 2 TAPAAEGAGAEVKASAEAK 21  
Db 16 SAAGAEGAGAPAAASAEPK 35

RESULT 6  
ABU17304  
ID ABU17304 standard; protein; 409 AA.

AC ABU17304;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #2831.

DE Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Acinetobacter baumannii.

PN WO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

PA (ELIT-) ELITRA PHARM INC.

PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Tamwick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

DR WPI; 2003-029926/02.

DR N-PSDB; ACA21174.

PS Claim 25; SEQ ID NO 45228; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of  
CC the 6213 antisense sequences given in the specification where expression  
CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
CC (1) a vector comprising a promoter operably linked to the nucleic acid  
CC encoding a polypeptide whose expression is inhibited by the antisense  
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
CC polypeptide or its fragment whose expression is inhibited by the  
CC antisense nucleic acid; (4) an antibody capable of specifically binding  
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
CC proliferation or the activity of a gene in an operon required for  
CC proliferation; (7) identifying a compound that influences the activity of  
CC the gene product or that has an activity against a biological pathway  
CC required for proliferation, or that inhibits cellular proliferation; (8)  
CC identifying a gene required for cellular proliferation or the biological  
CC pathway in which a proliferation-required gene or its gene product lies  
CC or a gene on which the test compound that inhibits proliferation of an  
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
CC compound's activity; (11) a culture comprising strains in which the gene  
CC product is overexpressed or underexpressed; (12) determining the extent  
CC to which each of the strains is present in a culture or collection of  
CC strains; or (13) identifying the target of a compound that inhibits the  
CC proliferation of an organism. The antisense nucleic acids are useful for  
CC identifying proteins or screening for homologous nucleic acids required  
CC for cellular proliferation to isolate candidate molecules for rational  
CC drug discovery programs, or for screening homologous nucleic acids  
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of  
CC the target prokaryotic essential genes. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 409 AA;

Query Match 44.3%; Score 51; DB 6; Length 409;  
Best Local Similarity 60.0%; Pred. No. 40;  
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKASAEAKQ 22

Db 269 APSGEGATPCIKLALAEAKQ 288

RESULT 7

ADE28111

ID ADE28111 standard; protein; 412 AA.

AC ADE28111;

DT 29-JAN-2004 (first entry)

DE Human NTRAN protein - SEQ ID 16.

DE human; neurotransmission-associated protein; NTRAN; cytostatic;

DE immunomodulator; immune disorder; cancer; gene therapy.

OS Homo sapiens.

PN WO2003051902-A1.

PD 26-JUN-2003.

PF 12-DEC-2002; 2002WO-US040059.

PR 14-DEC-2001; 2001US-0340798P.

PR 18-MAR-2002; 2002US-0365645P.

PR 25-MAR-2002; 2002US-0367662P.

PR 10-MAY-2002; 2002US-0379887P.

PR 31-MAY-2002; 2002US-0384639P.

XX (INCY-) INCYTE GENOMICS INC.

XX Baughn MR, Bhatia U, Blake JJ, Buzrill JD, Elliott VS;

PI Emerling EM, Forsythe IJ, Gietzen KJ, Gorvad AE, Griffin JA;

PI Hafalia AJA, Ho A, Jackson AA, Jiang X, Kable AE, Kearney L;

PI Khare R, Lee EA, Lee S, Lu DAM, Marquis JP, Lehr-Mason PM;

PI Ramkumar J, Richardson TW, Sprague WW, Tran UK, Chawla NK;

PI Warren BA, Yue H, Zheng W;

XX WPI; 2003-514037/48.

DR N-PSDB; ADE28133.

XX New human neurotransmission-associated proteins (NTRAN) polypeptide,

PT useful for preparing a composition for treating a disease associated with

PT decreased expression or overexpression of NTRAN e.g., cancer.

XX Claim 1; SEQ ID NO 16; 261pp; English.

XX The invention relates to a novel isolated human neurotransmission-

CC associated proteins (NTRAN) polypeptide. The polypeptide of the invention

CC demonstrates cytostatic and immunomodulator activities and may be useful

CC for preparing a composition for diagnosing or treating a disease or

CC condition associated with decreased expression or overexpression of

CC functional NTRAN including immune disorders or cancer, as well as during

CC gene therapy procedures. The current sequence is that of the human NTRAN

CC protein of the invention.

XX Sequence 412 AA;

SQ

Query Match 44.3%; Score 51; DB 7; Length 412;  
Best Local Similarity 60.0%; Pred. No. 40;  
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 TAPAEAGAGAEVKRASAEAK 21  
Db 16 SAAGAEGAGAPAAASAEKP 35

RESULT 8  
ADA33525  
ID ADA33525 standard; protein; 422 AA.

XX AC ADA33525;

XX DT 20-NOV-2003 (first entry)

XX DE Acinetobacter baumannii protein #686.

XX KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;  
XX KW plant biocontrol agent.

XX OS Acinetobacter baumannii.

XX FN US6562958-B1.

XX PD 13-MAY-2003.

XX PF 04-JUN-1999; 99US-00328352.

XX PR 09-JUN-1998; 98US-0088701P.

XX PA (GENO-) GENOME THERAPEUTICS CORP.

XX PI Breton G, Bush D;

XX DR WPI: 2003-576092/54.

XX DR N-PSDB; ADA29399.

XX PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents  
XX PT for diagnosing a bacterial disease, as components of antibacterial  
XX PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for  
XX PT plants.

XX PS Example; SEQ ID NO 4812; 328pp; English.

XX CC The invention relates to isolated Acinetobacter baumannii nucleic acids.  
XX CC The A. baumannii nucleic acids and polypeptides are useful as reagents  
XX CC for diagnosing a bacterial disease, as components of antibacterial  
XX CC vaccines, as targets for antibacterial drugs, to detect the presence of  
XX CC A. baumannii and other Acinetobacter species in a sample, in screening  
XX CC compounds for the ability to interfere with the A. baumannii life cycle  
XX CC or to inhibit A. baumannii infection, and as biocontrol agents for  
XX CC plants. The present sequence represents the amino acid sequence of an A.  
XX CC baumannii protein.

XX SQ Sequence 422 AA;

Query Match 44.3%; Score 51; DB 6; Length 422;  
Best Local Similarity 60.0%; Pred. No. 41;  
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAEAKQ 22  
Db 282 APSGEGATRCIKLALAEAKQ 301

RESULT 9  
ABU47184  
ID ABU47184 standard; protein; 488 AA.

XX AC ABU47184;

XX XX

DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #32711.

XX XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Salmomella typhimurium.

XX PN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX XX 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX PA (ELIT-) ELITRA PHARM INC.

XX XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX XX WPI: 2003-029926/02.

XX DR N-PSDB; ACA51054.

XX PS Claim 25; SEQ ID NO 75108; 1765pp; English.

XX CC The invention relates to an isolated nucleic acid comprising any one of  
XX CC the 6213 antisense sequences given in the specification where expression  
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:

XX CC (1) a vector comprising a promoter operably linked to the nucleic acid  
XX CC encoding a polypeptide whose expression is inhibited by the antisense  
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
XX CC polypeptide or its fragment whose expression is inhibited by the  
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding  
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
XX CC proliferation or the activity of a gene in an operon required for  
XX CC proliferation; (7) identifying a compound that influences the activity of  
XX CC the gene product or that has an activity against a biological pathway  
XX CC required for proliferation, or that inhibits cellular proliferation; (8)  
XX CC identifying a gene required for cellular proliferation or the biological  
XX CC pathway in which a proliferation-required gene or its gene product lies  
XX CC or a gene on which the test compound that inhibits proliferation of an  
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
XX CC compound's activity; (11) a culture comprising strains in which the gene  
XX CC product is overexpressed or underexpressed; (12) determining the extent  
XX CC to which each of the strains is present in a culture or collection of  
XX CC strains; or (13) identifying the target of a compound that inhibits the  
XX CC proliferation of an organism. The antisense nucleic acids are useful for  
XX CC identifying proteins or screening for homologous nucleic acids required  
XX CC for cellular proliferation to isolate candidate molecules for rational  
XX CC drug discovery programs, or for screening homologous nucleic acids  
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of  
XX CC the target prokaryotic essential genes. Note: The sequence data for this  
XX CC patent did not form part of the printed specification, but was obtained  
XX CC in electronic format directly from WIPO at  
XX CC ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 488 AA;

Query Match 44.3%; Score 51; DB 6; Length 488;  
Best Local Similarity 54.5%; Pred. No. 48;  
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAEAKQAF 24

Qy 2 TAPAAEGAGAEVKRAAEAK 21  
: | | | | | | | | | |  
Db 16 SAAGAEGAGAPAAAAAEPK 35

RESULT 11  
AAB94311  
ID AAB94311 standard: protein: 589 AA.

31 AEVSGGEAEVNQAVAAKEAF 52

RESULT 10  
AAB95205  
ID AAB95205 standard; protein: 589 AA.



```
PR 11-JAN-2000; 2000JP-00118776.
PR 17-FEB-2000; 2000US-0183322P.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T, Aburatani H;
PI Kodama T, Midorikawa Y;
XX
XX WPI; 2001-570287/64.
DR N-PSDB; AAI93885.
XX
XX New Stomach cancer-associated genes, useful as markers in blood tests for
PT screening for the early stages of the disease.
XX
XX Claim 1; Page 160-161; 242pp; Japanese.
XX
XX The invention relates to stomach cancer-expressed genes (AAI93842-
CC AAI93917) and the encoded proteins (AAM93967-AAM94039). The genes can be
CC used as markers in blood tests for screening for the early stages of the
CC disease. The proteins and peptides can be used as targets for screening
CC for compounds to treat the disease. They can also be used for predicting
CC micro-metastases. The gene can predict peritoneal dissemination
XX
XX Sequence 589 AA;
SQ
Query Match 44.3%; Score 51; DB 4; Length 589;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 2 TAPAAEGAGAEVKRAAEAK 21
DB 16 SAAGAEGAGAPAAAASAEPK 35
RESULT 15
ADE59377
ID ADE59377 standard; protein; 589 AA.
XX
AC ADE59377;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein NP_038466, SEQ ID NO 5271.
XX
XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
OS
XX WO2003016475-A2.
PN
XX 27-FEB-2003.
PD
XX
XX 14-AUG-2002; 2002WO-US025765.
PF
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
PA (PAB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
DR GENBANK; NP_038466.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
```

```
XX
XX Claim 1; Page: 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) or a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 589 AA;
SQ
Query Match 44.3%; Score 51; DB 7; Length 589;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 2 TAPAAEGAGAEVKRAAEAK 21
DB 16 SAAGAEGAGAPAAAASAEPK 35
Search completed: October 7, 2004, 17:34:01
Job time : 55.5 secs
```



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:35:42 ; Search time 84.5 Seconds  
(without alignments)  
91.399 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115  
Sequence: 1 GTAPAAEGAGAEVKRASAERAKQAF 24

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/FCI\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/FCIUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	100.0	24	US-10-180-753-2	Sequence 2, Appli
2	115	100.0	24	US-10-802-644-2	Sequence 2, Appli
3	53	46.1	406	US-10-437-963-147454	Sequence 147454,
4	53	46.1	487	US-10-282-122A-50399	Sequence 50399, A
5	52	45.2	116	US-10-437-963-178225	Sequence 178225,
6	52	45.2	177	US-10-767-701-34508	Sequence 34508, A
7	51	44.3	409	US-10-282-122A-45228	Sequence 45228, A
8	51	44.3	488	US-10-282-122A-75108	Sequence 75108, A
9	51	44.3	589	US-10-293-000-4	Sequence 2, Appli
10	51	44.3	595	US-10-293-000-2	Sequence 2, Appli
11	50.5	43.9	224	US-10-156-761-14054	Sequence 14054, A
12	50	43.5	298	US-10-425-114-47199	Sequence 47199, A
13	50	43.5	395	US-10-156-761-14429	Sequence 14429, A
14	49.5	43.0	74	US-09-864-408A-8324	Sequence 8324, Ap
15	49	42.6	143	US-10-437-963-165448	Sequence 165448,

16	49	42.6	292	16	US-10-437-963-170910	Sequence 170910,
17	49	42.6	298	12	US-10-425-114-56061	Sequence 56061, A
18	48.5	42.2	41	12	US-10-282-122A-70019	Sequence 70019, A
19	48.5	42.2	120	12	US-10-424-599-160168	Sequence 160168,
20	48.5	42.2	133	16	US-10-437-963-190078	Sequence 190078,
21	48.5	42.2	151	16	US-10-437-963-132303	Sequence 132303,
22	48.5	42.2	156	9	US-09-815-242-11996	Sequence 11996, A
23	48.5	42.2	156	12	US-10-282-122A-43550	Sequence 43550, A
24	48.5	42.2	156	12	US-10-282-122A-68349	Sequence 68349, A
25	48.5	42.2	156	14	US-10-127-032-107	Sequence 107, App
26	48.5	42.2	304	16	US-10-437-963-186384	Sequence 186384,
27	48.5	42.2	554	12	US-10-282-122A-50209	Sequence 50209, A
28	48	41.7	98	16	US-10-767-701-62081	Sequence 62081, A
29	48	41.7	152	16	US-10-437-963-108597	Sequence 108597,
30	48	41.7	562	16	US-10-437-963-114753	Sequence 114753,
31	48	41.7	618	9	US-09-925-300-1381	Sequence 1381, Ap
32	48	41.7	645	16	US-10-437-963-189302	Sequence 189302,
33	48	41.7	648	16	US-10-437-963-171650	Sequence 171650,
34	48	41.7	696	16	US-10-437-963-200901	Sequence 200901,
35	48	41.7	704	16	US-10-437-963-108152	Sequence 108152,
36	48	41.7	709	16	US-10-437-963-127877	Sequence 127877,
37	48	41.7	821	16	US-10-437-963-171696	Sequence 171696,
38	48	41.7	1036	12	US-10-425-114-67897	Sequence 67897, A
39	48	41.7	1140	16	US-10-437-963-171644	Sequence 171644,
40	48	41.7	1177	16	US-10-437-963-108155	Sequence 108155,
41	48	41.7	1184	16	US-10-437-963-171693	Sequence 171693,
42	48	41.7	1249	16	US-10-437-963-108149	Sequence 108149,
43	48	41.7	1274	16	US-10-437-963-108146	Sequence 108146,
44	48	41.7	1387	14	US-10-156-761-13000	Sequence 13000, A
45	48	41.7	1527	16	US-10-437-963-200899	Sequence 200899,

ALIGNMENTS

RESULT 1  
US-10-180-753-2  
; Sequence 2, Application US/10180753  
; Publication No. US20030013652A1  
; GENERAL INFORMATION:  
; APPLICANT: Martin, Linda  
; APPLICANT: Adler, Kenneth  
; APPLICANT: Li, Yuehua  
; TITLE OF INVENTION: BLOCKING PEPTIDE FOR INFLAMMATORY CELL SECRETION  
; FILE REFERENCE: 5051.574  
; CURRENT APPLICATION NUMBER: US/10/180,753  
; CURRENT FILING DATE: 2002-06-26  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic peptide  
; NAME/KEY: MOD RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: N-terminal myristoylation  
US-10-180-753-2

Query Match 100.0%; Score 115; DB 12; Length 24;  
Best Local Similarity 100.0%; Pred. No. 4.3e-09;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVKRASAERAKQAF 24  
|||||  
Db 1 GTAPAAEGAGAEVKRASAERAKQAF 24  
|||||

RESULT 2  
US-10-802-644-2  
; Sequence 2, Application US/10802644

Publication No. US20040180836A1

GENERAL INFORMATION:  
APPLICANT: Martin, Linda  
APPLICANT: Li, Yuehua  
TITLE OF INVENTION: BLOCKING PEPTIDE FOR INFLAMMATORY CELL SECRETION  
FILE REFERENCE: 5051.574  
CURRENT APPLICATION NUMBER: US/10/802,644  
CURRENT FILING DATE: 2004-03-17  
PRIOR APPLICATION NUMBER: US/10/180,753  
PRIOR FILING DATE: 2002-06-26  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 2  
LENGTH: 24  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: Synthetic peptide  
NAME/KEY: MOD RES  
LOCATION: (1)-(1)  
OTHER INFORMATION: N-terminal myristoylation  
US-10-802-644-2

Query Match 100.0%; Score 115; DB 16; Length 24;  
Best Local Similarity 100.0%; Pred. No. 4.3e-09;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVVKRAAEAKQAF 24  
|||||  
DB 1 GTAPAAEGAGAEVVKRAAEAKQAF 24

## RESULT 3

US-10-437-963-147454  
Sequence 147454, Application US/10437963  
Publication No. US20040123343A1

GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963  
CURRENT FILING DATE: 2003-05-14  
NUMBER OF SEQ ID NOS: 204966  
SEQ ID NO 147454  
LENGTH: 406  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT4530\_47981C.1.pep  
US-10-437-963-147454

Query Match 46.1%; Score 53; DB 16; Length 406;  
Best Local Similarity 52.6%; Pred. No. 35;  
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVVKRAAE 19  
|||||  
DB 6 GGTPEEGGGADVEKAAAE 24

## RESULT 4

US-10-282-122A-50399  
Sequence 50399, Application US/10282122A

Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Haselbeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
APPLICANT: Xu, H.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA\_034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
CURRENT FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 50399  
LENGTH: 487  
TYPE: PRT  
ORGANISM: Burkholderia mallei  
US-10-282-122A-50399

Query Match 46.1%; Score 53; DB 12; Length 487;  
Best Local Similarity 53.6%; Pred. No. 42;  
Matches 15; Conservative 4; Mismatches 5; Indels 4; Gaps 2;

QY 1 GTA-PAAEGAGA--EVKRAAEAKQAF 24  
|||||  
DB 24 GIGEPVWEGAGASADDERAVASARRAF 51

## RESULT 5

US-10-437-963-178225  
Sequence 178225, Application US/10437963  
Publication No. US20040123343A1

GENERAL INFORMATION:  
APPLICANT: La Rosa, Thomas J.  
APPLICANT: Kovalic, David K.  
APPLICANT: Zhou, Yihua  
APPLICANT: Cao, Yongwei  
APPLICANT: Wu, Wei  
APPLICANT: Boukharov, Andrey A.  
APPLICANT: Barbazuk, Brad  
APPLICANT: Li, Ping  
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with  
FILE REFERENCE: 38-21(53221)B  
CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 178225  
; LENGTH: 116  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(116)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_75801C.1.pep  
US-10-437-963-178225

Query Match 45.2%; Score 52; DB 16; Length 116;  
Best Local Similarity 55.0%; Pred. No. 13;  
Matches 11; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVKRASAEA 20  
Db 96 GDRPAEGSGEAAALAA 115

## RESULT 6

US-10-767-701-34508  
; Sequence 34508, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53535)B  
; CURRENT APPLICATION NUMBER: US/10767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 34508  
; LENGTH: 177  
; TYPE: PRT  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(177)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C4969\_1.pep  
US-10-767-701-34508

Query Match 45.2%; Score 52; DB 16; Length 177;  
Best Local Similarity 52.4%; Pred. No. 20;  
Matches 11; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAEAKQA 23  
Db 101 SPATGGAAKKORASIMAKQS 121

## RESULT 7

US-10-282-122A-45228  
; Sequence 45228, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 45228  
; LENGTH: 409  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-10-282-122A-45228

Query Match 44.3%; Score 51; DB 12; Length 409;  
Best Local Similarity 60.0%; Pred. No. 67;  
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAEAKQ 22  
Db 269 APSGEGATRCIKLALAEAKQ 288

## RESULT 8

US-10-282-122A-75108  
; Sequence 75108, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06

```
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 75108
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
US-10-282-122A-75108

Query Match 44.3%; Score 51; DB 12; Length 488;
Best Local Similarity 54.5%; Pred. No. 80;
Matches 12; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 3 AAPAEGAGAEVKRASAEAKQAF 24
DB 31 AEVASGGAEVNVQAVAAKEAF 52

RESULT 9
US-10-293-000-4
; Sequence 4, Application US/10293000
; Publication No. US20030175278A1
; GENERAL INFORMATION:
; APPLICANT: Monteiro, Mervyn J.
; APPLICANT: Mah, Alex L.
; APPLICANT: Perry, George
; APPLICANT: Smith, Mark A.
; TITLE OF INVENTION: UBQUILIN, A PRESENILIN INTERACTOR AND METHODS OF USING SAME
; FILE REFERENCE: 4115-175
; CURRENT APPLICATION NUMBER: US/10/293,000
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/338,549
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 4
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-000-4

Query Match 44.3%; Score 51; DB 14; Length 589;
Best Local Similarity 60.0%; Pred. No. 98;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 TAPAAEGAGAEVKRASAEAK 21
DB 16 SAAGAEGAGAPAAASAEPK 35

RESULT 10
US-10-293-000-2
; Sequence 2, Application US/10293000
; Publication No. US20030175278A1
; GENERAL INFORMATION:
; APPLICANT: Monteiro, Mervyn J.
; APPLICANT: Mah, Alex L.
; APPLICANT: Perry, George
; APPLICANT: Smith, Mark A.
; TITLE OF INVENTION: UBQUILIN, A PRESENILIN INTERACTOR AND METHODS OF USING SAME
; FILE REFERENCE: 4115-175
```

```
; CURRENT APPLICATION NUMBER: US/10/293,000
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: US 60/338,549
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 2
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-000-2

Query Match 44.3%; Score 51; DB 14; Length 595;
Best Local Similarity 60.0%; Pred. No. 99;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 TAPAAEGAGAEVKRASAEAK 21
DB 16 SAAGAEGAGAPAAASAEPK 35

RESULT 11
US-10-156-761-14054
; Sequence 14054, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 14054
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-14054

Query Match 43.9%; Score 50.5; DB 14; Length 224;
Best Local Similarity 40.6%; Pred. No. 42;
Matches 13; Conservative 4; Mismatches 6; Indels 9; Gaps 1;

QY 1 GTAPAAEG-----AGAEVKRASAEAKQA 23
DB 124 GSAPAAAGTGAGSGSSAGSATTESAAKQA 155

RESULT 12
US-10-425-114-47199
; Sequence 47199, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
```

; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 47199  
; LENGTH: 298  
; TYPE: PRT  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3079-043-H2\_FLI.pep  
US-10-425-114-47199

Query Match 43.5%; Score 50; DB 12; Length 298;  
Best Local Similarity 50.0%; Pred. No. 66;  
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 APAEAGAGAEVVKASAEAKQ 22  
Db 182 APAEAGAGAAVQGGGERRR 201

RESULT 13  
US-10-156-761-14429  
; Sequence 14429, Application US/10156761  
; Publication No. US20030119018A1  
; GENERAL INFORMATION:  
; APPLICANT: OMURA, SATOSHI  
; APPLICANT: IKEDA, HARUO  
; APPLICANT: ISHIKAWA, JUN  
; APPLICANT: HORIKAWA, HIROSHI  
; APPLICANT: SHIBA, TADAYOSHI  
; APPLICANT: SAKAKI, YOSHIYUKI  
; APPLICANT: HATTORI, MASAHIRA  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-262  
; CURRENT APPLICATION NUMBER: US/10/156,761  
; CURRENT FILING DATE: 2002-05-29  
; PRIOR APPLICATION NUMBER: JP 2001-204089  
; PRIOR FILING DATE: 2001-05-30  
; PRIOR APPLICATION NUMBER: JP 2001-272697  
; PRIOR FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 15109  
; SEQ ID NO 14429  
; LENGTH: 395  
; TYPE: PRT  
; ORGANISM: Streptomyces avermitilis  
US-10-156-761-14429

Query Match 43.5%; Score 50; DB 14; Length 395;  
Best Local Similarity 60.0%; Pred. No. 89;  
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVVKASAEA 20  
Db 116 GTGPAVEQAVALVERAAKEA 135

RESULT 14  
US-09-864-408A-8324  
; Sequence 8324, Application US/09864408A  
; Publication No. US20040009474A1  
; GENERAL INFORMATION:  
; APPLICANT: Leach, Martin D.  
; APPLICANT: Shinkets, Richard A.  
; TITLE OF INVENTION: No. US20040009474A1 Human Polynucleotides and Polypeptides Encod  
; FILE REFERENCE: 21402-012  
; CURRENT APPLICATION NUMBER: US/09/864,408A  
; CURRENT FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 60/206,690  
; PRIOR FILING DATE: 2000-05-24  
; NUMBER OF SEQ ID NOS: 9068  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8324  
; LENGTH: 74  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-09-864-408A-8324

Query Match 43.0%; Score 49.5; DB 11; Length 74;  
Best Local Similarity 41.9%; Pred. No. 18;  
Matches 13; Conservative 5; Mismatches 4; Indels 9; Gaps 1;

QY 2 TAPAAEAGA-----GAEVKRAAEAKOA 23  
Db 33 TAPVEGAVVAAMVMASTGALELVVAAAEAKA 63

RESULT 15  
US-10-437-963-165448  
; Sequence 165448, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 165448  
; LENGTH: 143  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(143)  
; OTHER INFORMATION: unsure at all xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_64253C.1.pep  
US-10-437-963-165448

Query Match 42.6%; Score 49; DB 16; Length 143;  
Best Local Similarity 43.8%; Pred. No. 42;  
Matches 14; Conservative 0; Mismatches 8; Indels 10; Gaps 1;

QY 1 GTAPAAEGAGAEVVKRA-----SAEAKQ 22  
Db 93 GDRPAAEGGGVAATRAATAVAKRLGGSAAKQ 124

Search completed: October 7, 2004, 17:50:36  
Job time : 86.5 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:32:03 ; Search time 15 Seconds  
(without alignments)  
82.602 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAAEGAGAEVKRAAEAKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B-COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTUS-COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	45.2	211	4	US-09-252-991A-24706
2	52	45.2	407	4	US-09-252-991A-31200
3	52	45.2	453	4	US-09-252-991A-24290
4	51	44.3	422	4	US-09-328-352-4812
5	51	44.3	589	4	US-09-665-479A-4
6	50	43.5	259	4	US-09-252-991A-23293
7	50	43.5	492	4	US-09-489-039A-9023
8	49	42.6	173	4	US-09-252-991A-27742
9	49	42.6	418	4	US-09-328-352-5700
10	48.5	42.2	163	4	US-09-252-991A-21344
11	48.5	42.2	197	4	US-09-252-991A-22076
12	48	41.7	402	4	US-09-252-991A-29857
13	47.5	41.3	181	4	US-09-328-352-4739
14	47	40.9	31	4	US-09-690-454-159
15	47	40.9	125	4	US-09-690-454-155
16	47	40.9	143	1	US-08-253-155A-46
17	47	40.9	143	4	US-09-489-039A-12453
18	47	40.9	165	4	US-09-252-991A-17356
19	47	40.9	168	4	US-09-489-039A-7835
20	47	40.9	218	3	US-08-675-885-5
21	47	40.9	256	4	US-09-252-991A-26244
22	47	40.9	317	4	US-09-328-352-6383
23	47	40.9	329	4	US-09-252-991A-21954
24	47	40.9	378	3	US-09-032-372-11
25	47	40.9	378	3	US-08-853-733B-2
26	47	40.9	378	3	US-08-675-885-7
27	47	40.9	498	4	US-09-489-039A-7569

28 47 40.9 610 4 US-09-252-991A-20299 Sequence 20299, A  
29 46 40.0 136 4 US-09-732-210-814 Sequence 814, App  
30 46 40.0 136 4 US-09-732-210-815 Sequence 815, App  
31 46 40.0 219 2 US-08-557-309B-54 Sequence 54, Appl  
32 46 40.0 249 4 US-09-252-991A-30556 Sequence 30556, A  
33 46 40.0 284 3 US-08-564-164A-2 Sequence 2, Appl  
34 46 40.0 442 3 US-08-834-306-52 Sequence 52, Appl  
35 46 40.0 442 3 US-08-993-674A-52 Sequence 52, Appl  
36 46 40.0 442 4 US-09-256-976-52 Sequence 52, Appl  
37 46 40.0 500 4 US-09-354-123-6 Sequence 1, Appl  
38 46 40.0 571 2 US-08-796-414B-1 Sequence 1, Appl  
39 46 40.0 584 4 US-09-198-452A-352 Sequence 352, App  
40 46 40.0 612 4 US-09-252-991A-25727 Sequence 25727, A  
41 46 40.0 767 4 US-09-252-991A-28484 Sequence 28484, A  
42 45.5 39.6 388 4 US-09-252-991A-30608 Sequence 30608, A  
43 45 39.1 84 4 US-09-489-039A-10806 Sequence 10806, A  
44 45 39.1 444 3 US-09-252-292C-27 Sequence 27, Appl  
45 45 39.1 472 4 US-09-252-991A-29165 Sequence 29165, A

#### ALIGNMENTS

RESULT 1  
US-09-252-991A-24706  
; Sequence 24706, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 24706  
; LENGTH: 211  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-24706

Query Match 45.2%; Score 52; DB 4; Length 211;  
Best Local Similarity 52.2%; Pred. No. 3.7;  
Matches 12; Conservative 2; Mismatches 9; Indels 0; Gaps 0;  
QY 1 GTAPAAEGAGAEVKRAAEAKQA 23  
Db 79 GSAPATRRAGAAADRAAARAAA 101

RESULT 2  
US-09-252-991A-31200  
; Sequence 31200, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 31200  
; LENGTH: 407  
; TYPE: PRT

```
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31200

Query Match          45.2%; Score 52; DB 4; Length 407;
Best Local Similarity 59.1%; Pred. No. 7.5;
Matches 13; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY      2 TAPAAEGAGAEVVKRASAEAKQ 23
      :|||:|||||:|||||:
Db      220 SATAAGGPGAEQARAEREAGAA 241

RESULT 3
US-09-252-991A-24290
; Sequence 24290, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24290
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24290

Query Match          45.2%; Score 52; DB 4; Length 453;
Best Local Similarity 64.7%; Pred. No. 8.5;
Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY      3 APAAEGAGAEVVKRASAE 19
      :|||:|||||:|||||:
Db      421 APAAEGGAGAEQQAQAQAD 437

RESULT 4
US-09-328-352-4812
; Sequence 4812, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4812
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4812

Query Match          44.3%; Score 51; DB 4; Length 422;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      3 APAAEGAGAEVVKRASAEAKQ 22
      :|||:|||||:|||||:
Db      282 APSGEGATRCIKLALAEAKQ 301

RESULT 5
US-09-665-479A-4
; Sequence 4, Application US/09665479A
; Patent No. 6673570
; GENERAL INFORMATION:
; APPLICANT: Itoh, Fumiko
; APPLICANT: Itoh, Susumu
; APPLICANT: Heldin, Carl-Henrik
; APPLICANT: ten-Dijke, Peter
; TITLE OF INVENTION: SMAD ASSOCIATING POLYPEPTIDES
; FILE REFERENCE: L00461.70096.US
; CURRENT APPLICATION NUMBER: US/09/665,479A
; CURRENT FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US 60/154,846
; PRIOR FILING DATE: 1999-09-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-665-479A-4

Query Match          44.3%; Score 51; DB 4; Length 589;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 12; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY      2 TAPAAEGAGAEVVKRASAEAK 21
      :|||:|||||:|||||:
Db      16 SAAGAEGAGAPAAAAASAEPK 35

RESULT 6
US-09-252-991A-23293
; Sequence 23293, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23293
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23293

Query Match          43.5%; Score 50; DB 4; Length 259;
Best Local Similarity 47.8%; Pred. No. 9;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY      1 GTAPAAEGAGAEVVKRASAEAKQA 23
      :|||:|||||:|||||:
Db      152 GNAPAGQGAGAEERSRIARHLEA 174

RESULT 7
US-09-489-039A-9023
; Sequence 9023, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
```



26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528  
 529  
 530  
 531  
 532  
 533  
 534  
 535  
 536  
 537  
 538  
 539  
 540  
 541  
 542  
 543  
 544  
 545  
 546

2

; CURRENT APPLICATION NUMBER: US/09/252,991A

```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29857
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29857

Query Match          41.7%; Score 48; DB 4; Length 402;
Best Local Similarity 47.8%; Pred. No. 28;
Matches 11; Conservative 3; Mismatches 9; Indels 9; Gaps 0;

Qy 1 GTAPAAEGAGAEVKRASAERKQA 23
Db 26 GGDPSVAGGAGVQRAPGEAVRA 48

RESULT 13
US-09-328-352-4739
; Sequence 4739, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4739
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4739

Query Match          41.3%; Score 47.5; DB 4; Length 181;
Best Local Similarity 48.0%; Pred. No. 14;
Matches 12; Conservative 3; Mismatches 5; Indels 5; Gaps 1;

Qy 5 AAEAGAGAEVKRAS-----AEAKQAF 24
Db 152 AAEAGGAIAKKREDVHRMAEANKAF 176

RESULT 14
US-09-690-454-159
; Sequence 159, Application US/09690454
; Patent No. 6531447
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: PZ006P1
; CURRENT APPLICATION NUMBER: US/09/690,454
; CURRENT FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 155
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-690-454-155

Query Match          40.9%; Score 47; DB 4; Length 125;
Best Local Similarity 76.9%; Pred. No. 11;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 AEGAGAEVKRASA 18
Db 6 AEGAGAEVETATA 18
```

```
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 159
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-690-454-159

Query Match          40.9%; Score 47; DB 4; Length 31;
Best Local Similarity 76.9%; Pred. No. 2.5;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 AEGAGAEVKRASA 18
Db 6 AEGAGAEVETATA 18

RESULT 15
US-09-690-454-155
; Sequence 155, Application US/09690454
; Patent No. 6531447
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: PZ006P1
; CURRENT APPLICATION NUMBER: US/09/690,454
; CURRENT FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 155
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (30)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-690-454-155

Query Match          40.9%; Score 47; DB 4; Length 125;
Best Local Similarity 76.9%; Pred. No. 11;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 AEGAGAEVKRASA 18
Db 6 AEGAGAEVETATA 18
```

Tue Oct 12 05:37:43 2004

us-10-802-644-2.open.ra1

Page 5

Db 43 AEGAGAEVETATA 55

Search completed: October 7, 2004, 17:36:36  
Job time : 17 secs

**This Page Blank (uspto)**

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	27.5	23.9	16	2	PC4371	telomeric and tetr	
2	27	23.5	22	2	T15870	hypothetical prote	
3	26	22.6	10	2	PS0209	24K protein 4407 -	
4	26	22.6	15	2	A49177	22K protein p1, mi	
5	26	22.6	20	2	PQ0688	photosystem I 14.0	
6	26	22.6	20	2	T25653	hypothetical prote	
7	26	22.6	22	2	S68900	xanthine dehydroge	
8	26	22.6	22	2	B48186	ATP synthase beta-	
9	25	21.7	17	2	C30221	histone H2A.8 - ch	
10	25	21.7	19	2	A41299	T-cell receptor al	
11	25	21.7	20	2	PQ0687	photosystem I 14.1	
12	25	21.7	21	2	S65611	tubulin beta chain	
13	25	21.7	21	2	PC7043	ubiquitin carboxyl	
14	25	21.7	22	2	H30608	Ig kappa chain V-I	
15	25	21.7	22	2	PQ0143	polygalacturonase	
16	25	21.7	24	1	BMTD	bovinin - Bombina	
17	24	20.9	14	2	B39111	Ig heavy chain V r	
18	24	20.9	14	2	PH1325	Ig heavy chain DJ	
19	24	20.9	14	2	PH0755	T-cell receptor be	
20	24	20.9	15	2	PQ0692	photosystem I 18.5	
21	24	20.9	15	2	PT0037	light harvesting c	
22	24	20.9	16	2	A27803	myosin light chain	
23	24	20.9	17	2	E29501	fibrinopeptide A -	
24	24	20.9	18	2	S45373	translation elonga	
25	24	20.9	19	2	B43978	urotensin I precur	
26	24	20.9	20	2	A45806	T-cell receptor be	
27	24	20.9	20	2	S06149	photosystem I chai	
28	24	20.9	20	2	S68617	histone H2A - sea	
29	24	20.9	20	2	B60801	acrosome stabiliz	

## RESULT 3

PS0209  
 24K protein 4407 - rice (strain Nihonbare) (fragment)  
 C;Species: Oryza sativa (rice)  
 C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 11-Apr-1995  
 C;Accession: PS0209  
 R;Tsugita, A.; Kamo, M.  
 Submitted to JIPID, April 1993  
 A;Reference number: PS0209  
 A;Accession: PS0209  
 A;Molecule type: protein  
 A;Residues: 1-10 <TSU>  
 A;Experimental source: callus  
 C;Comment: molecular weight 24K, pI 4.6.

Query Match 22.6%; Score 26; DB 2; Length 10;  
 Best Local Similarity 75.0%; Pred. No. 1.3e+03;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 15 AERPGEAA 22  
 :|||:|  
 Db 3 AEVPAEAA 10

## RESULT 4

A49177  
 22K protein p1, mitochondrial - Crithidia fasciculata (fragment)  
 C;Species: Crithidia fasciculata  
 C;Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 19-May-1994  
 C;Accession: A49177; PH1826  
 R;Tittawella, I.

Exp. Cell Res. 206, 143-151, 1993  
 A;Title: Identification of DNA-binding proteins in the parasitic protozoan Crithidia fasciculata  
 A;Reference number: A49177; PMID:8482355  
 A;Accession: A49177  
 A;Molecule type: protein  
 A;Residues: 1-15 <TIT>  
 A;Note: sequence extracted from NCBI backbone (NCBIP:1311148)  
 C;Comment: This protein binds with mitochondrial DNA.  
 C;Keywords: DNA binding; mitochondrion

Query Match 22.6%; Score 26; DB 2; Length 15;  
 Best Local Similarity 46.2%; Pred. No. 1.9e+03;  
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 12 EAARPPGEAA 24  
 :|||:|  
 Db 1 DAPASAPKAAA 13

## RESULT 5

PQ0688  
 Photosystem I 14.0K E4 chain - common tobacco (fragment)  
 C;Species: Nicotiana tabacum (common tobacco)  
 C;Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 17-Mar-1999  
 C;Accession: PQ0688  
 R;Obokata, J.; Mikami, K.; Hayashida, N.; Nakamura, M.; Sugiyura, M.  
 Plant Physiol. 102, 1259-1267, 1993  
 A;Title: Molecular heterogeneity of photosystem I. psad, psae, psaf, psah and psal are a  
 A;Reference number: PQ0667; PMID:8278548  
 A;Accession: PQ0688  
 A;Molecule type: protein  
 A;Residues: 1-20 <OBO>  
 C;Keywords: chloroplast; photosynthesis; photosystem I

Query Match 22.6%; Score 26; DB 2; Length 20;  
 Best Local Similarity 40.0%; Pred. No. 2.4e+03;  
 Matches 8; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

Qy 5 SKTAKGEAARPPGEAA 24  
 :|||:|  
 Db 1 AEEAAPATATAEGEAPP 20

## RESULT 6

T25653  
 hypothetical protein C47C12.5 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
 C;Accession: T25653  
 R;Connell, M.  
 Submitted to the EMBL Data Library, August 1996  
 A;Description: The sequence of C. elegans cosmid C47C12.  
 A;Reference number: Z20062  
 A;Accession: T25653  
 A;Status: preliminary; translated from GB/EMBL/DBDJ  
 A;Molecule type: DNA  
 A;Residues: 1-22 <CON>

A;Cross-references: EMBL:U67951; PIDN:AAB07573.1; GSPDB:GNC0028; CESP:C47C12.5  
 A;Experimental source: strain Bristol N2; clone C47C12  
 C;Genetics:  
 A;Gene: CESP:C47C12.5  
 A;Map position: X  
 A;Introns: 21/1

Query Match 22.6%; Score 26; DB 2; Length 22;  
 Best Local Similarity 50.0%; Pred. No. 2.6e+03;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 QFSKTAAKGE 12  
 |||:|  
 Db 12 QFSRISLQNE 21

## RESULT 7

S68900  
 xanthine dehydrogenase (EC 1.1.1.204) 82.4K chain - Veillonella atypica (fragment)  
 C;Species: Veillonella atypica  
 C;Date: 23-Jul-1997 #sequence\_revision 29-Aug-1997 #text\_change 29-Aug-1997  
 C;Accession: S68900  
 R;Grewer, L.; Meyer, O.  
 Eur. J. Biochem. 238, 862-866, 1996  
 A;Title: Characterization of xanthine dehydrogenase from the anaerobic bacterium Veillonella  
 A;Reference number: S68900; MUID:96300255; PMID:8706691

A;Accession: S68900  
 A;Molecule type: protein  
 A;Residues: 1-22 <GRE>  
 A;Experimental source: DSM 1399  
 C;Keywords: 2Fe-2S; FAD; flavoprotein; heterotrimer; iron-sulfur protein; metalloprotein

Query Match 22.6%; Score 26; DB 2; Length 22;  
 Best Local Similarity 45.5%; Pred. No. 2.6e+03;  
 Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAGK 11  
 :|||:|  
 Db 5 GKSVDKVDAGK 15

## RESULT 8

B48186  
 ATP synthase beta-2 chain - wood tobacco (fragment)  
 C;Species: Nicotiana sylvestris (wood tobacco)  
 C;Date: 16-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 23-Mar-1995  
 C;Accession: B48186  
 R;De Paape, R.; Forchioni, A.; Chetrit, P.; Vedel, F.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 5934-5938, 1993  
 A;Title: Specific mitochondrial proteins in pollen: presence of an additional ATP synthase  
 A;Reference number: A48186; MUID:93317598; PMID:8327463

A;Accession: B48186  
 A;Status: preliminary  
 A;Molecule type: protein  
 A;Residues: 1-23 <DEI>  
 A;Experimental source: pollen  
 A;Note: sequence extracted from NCBI backbone (NCBIP:134868)

```

Query Match      22.6%; Score 26; DB 2; Length 23;
Best Local Similarity 83.3%; Pred. No. 2.7e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 13 AAAERP 18
DB 4 AAAEXP 9

RESULT 9
histone H2A.8 - chicken (fragment)
C:Species: Gallus gallus (chicken)
C>Date: 02-Nov-1989 #sequence_revision 02-Nov-1989 #text_change 30-Sep-1993
C:Accession: C30221
R:Challoner, P.B.; Moss, S.B.; Groudine, M.
Mol. Cell. Biol. 9, 902-913, 1989
A>Title: Expression of replication-dependent histone genes in avian spermatids involves
A:Reference number: A30221; MUID:89261754; PMID:2471062
A:Accession: C30221
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-17 <CHA>

Query Match      21.7%; Score 25; DB 2; Length 17;
Best Local Similarity 46.2%; Pred. No. 2.9e+03;
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGEA 13
DB 5 GKQGGKARAKA 17

RESULT 10
A41299
T-cell receptor alpha chain precursor V region (17.2) - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 23-Jul-1999
C:Accession: A41299
R:Uematsu, Y.; Wege, H.; Straus, A.; Ott, M.; Bannwarth, W.; Lanchbury, J.; Panayi, G.;
Proc. Natl. Acad. Sci. U.S.A. 88, 8534-8538, 1991
A>Title: The T-cell-receptor repertoire in the synovial fluid of a patient with rheumatoid
A:Reference number: A41299; MUID:92020887; PMID:1656449
A:Accession: A41299
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-19 <UEM>
A:Cross-references: GB:S57433; NID:g236318; PIDN:AAE19956.1; PID:g236319
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: T-cell receptor

Query Match      21.7%; Score 25; DB 2; Length 19;
Best Local Similarity 57.1%; Pred. No. 3.2e+03;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 17 RPGEAAV 23
DB 9 QPGDSAV 15

RESULT 11
PQ0687
photosystem I 14.1K E3 chain - common tobacco (fragment)
C:Species: Nicotiana tabacum (common tobacco)
C>Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 17-Mar-1999
C:Accession: PQ0687; PQ0675
R:Obokata, J.; Mikami, K.; Hayashida, N.; Nakamura, M.; Sugiyura, M.
Plant Physiol. 102, 1259-1267, 1993
A>Title: Molecular heterogeneity of photosystem I. psaD, psaE, psaF, psaH and psaL are a
A:Reference number: PQ0667; MUID:94105345; PMID:8278548
A:Accession: PQ0687
A:Molecule type: protein
A:Residues: 1-20 <OBO>

```

```

A>Note: this peptide was designated photosystem I 14.1K E3 chain
A:Accession: PQ0675
A:Molecule type: protein
A:Residues: 2-11 <OB2>
A>Note: this peptide was designated photosystem I 14.0K E4 chain
C:Keywords: chloroplast; photosynthesis; photosystem I; thylakoid

Query Match      21.7%; Score 25; DB 2; Length 20;
Best Local Similarity 41.2%; Pred. No. 3.3e+03;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 5 SKTAAGGAAARPCGEA 21
DB 2 AEEAAPPAATATAGEA 18

RESULT 12
S65611
tubulin beta chain - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
C:Accession: S65611
R:Ruediger, M.; Plessmann, U.; Ruediger, A.H.; Weber, K.
FEBS Lett. 364, 147-151, 1995
A>Title: Beta tubulin of bull sperm is polyglycylated.
A:Reference number: S65611; MUID:95269788; PMID:7750559
A:Accession: S65611
A:Molecule type: protein
A:Residues: 1-21 <RUE>
C:Complex: heterodimer; alpha and beta chain
C:Superfamily: tubulin
C:Keywords: GTP binding; heterodimer; microtubule

Query Match      21.7%; Score 25; DB 2; Length 21;
Best Local Similarity 55.6%; Pred. No. 3.5e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 12 EAAARPCGE 20
DB 12 DATAEEGE 20

RESULT 13
PC7043
ubiquitin carboxyl-terminal hydrolase (EC 3.1.1.-) homolog - yeast (Pichia anomala) (frag
N:Alternate names: ALX1 protein
C:Species: Pichia anomala, Candida pelliculosa
C>Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 11-May-2000
C:Accession: PC7043
R:Sakajo, S.; Minagawa, N.; Yoshimoto, A.
Biosci. Biotechnol. Biochem. 63, 1889-1894, 1999
A>Title: Structure and regulatory expression of a single copy alternative oxidase gene fr
A:Reference number: JC7145; MUID:20101471; PMID:10635554
A:Accession: PC7043
A:Molecule type: DNA
A:Residues: 1-21 <SAK>
A:Cross-references: DDBJ:AB026726
C:Genetics:
A:Gene: alx1
C:Keywords: hydrolase

Query Match      21.7%; Score 25; DB 2; Length 21;
Best Local Similarity 37.5%; Pred. No. 3.5e+03;
Matches 6; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 5 SKTAAGGAAARPCGE 20
DB 3 AKEAKVGSVEDKEGE 18

RESULT 14
H30608
Ig kappa chain V-III region (Ste) - human (fragment)

```

C;Species: Homo sapiens (man)  
C;Date: 29-Jun-1989 #sequence\_revision 29-Jun-1989 #text\_change 30-May-1997  
C;Accession: H30608  
R;Goni, F.R.; Chen, P.P.; McGinnis, D.; Arjonilla, M.L.; Fernandez, J.; Carson, D.; Sold  
J. Immunol. 142, 3158-3163, 1989  
A;Title: Structural and idiotypic characterization of the L chains of human IGM autoantib  
A;Reference number: A30601; MUID:89215279; PMID:2496160  
A;Accession: H30608

A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-22 <GON>  
C;Superfamily: immunoglobulin V region; immunoglobulin homology  
C;Keywords: heterotetramer; immunoglobulin

Query Match 21.7%; Score 25; DB 2; Length 22;  
Best Local Similarity 71.4%; Pred. No. 3.6e+03;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 18 PGEAAVA 24  
||| |  
Db 15 PGERATA 21

## RESULT 15

PQ0143  
polygalacturonase (EC 3.2.1.15) P26 - evening primrose (fragment)  
C;Species: Oenothera organensis (evening primrose)  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 06-Dec-1996  
C;Accession: PQ0143

R;Brown, S.M.; Crouch, M.L.  
Plant Cell 2, 263-274, 1990  
A;Title: Characterization of a gene family abundantly expressed in Oenothera organensis  
A;Reference number: JQ0992; MUID:93005658; PMID:2152116  
A;Accession: PQ0143

A;Molecule type: mRNA  
A;Residues: 1-22 <BRO>  
A;Experimental source: pollen  
C;Comment: This protein is specifically translated in the pollens.  
C;Comment: This protein functions by depolymerizing pectin in the cell walls of the pist  
V growing tube.  
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 21.7%; Score 25; DB 2; Length 22;  
Best Local Similarity 50.0%; Pred. No. 3.6e+03;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 SKTAAKGEAA 14  
:| ||| |  
Db 12 TKYAKGKGGA 21

Search completed: October 7, 2004, 17:57:24  
Job time : 26 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:36:42 ; Search time 14 seconds  
(without alignments)  
89.263 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAGGAAARPGCAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 1550

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	27.0	21	1 MISG MISAN	P81474 misgurnus a
2	27	23.5	15	1 UC23 MAIZE	P80629 zea mays (m
3	27	23.5	12	1 YOR8 CAEBL	Q09290 caenorhabdi
4	26	22.6	18	1 FIBA LAMGL	P14454 lama glama
5	25	21.7	24	1 BOWN BOWVA	P01505 bombina var
6	24	20.9	14	1 MCR2 METTM	P58816 methanobact
7	24	20.9	16	1 FIBA-TAPTE	P14536 capirus ter
8	24	20.9	17	1 FIBA-PIG	P14460 sus scrofa
9	24	20.9	23	1 IBP2 PIG	P24853 sus scrofa
10	23	20.0	16	1 DBH3 RHILE	P80605 rhizobium l
11	23	20.0	19	1 FIBA-BUBBU	P14442 bubalus bub
12	23	20.0	20	1 COXQ-THUOB	P80983 thunnus obe
13	23	20.0	20	1 PORC-METTM	P80902 methanobact
14	23	20.0	24	1 FK84 RAT	Q9qvc8 rattus norv
15	22	19.1	18	1 FIBA-CAMDR	P14444 camelus dro
16	22	19.1	20	1 CRP MUSCA	P19094 mustelus ca
17	22	19.1	23	1 VG22-BPT2	P21596 bacterioph
18	22	19.1	23	1 VG22-BPT2	P21596 bacterioph
19	22	19.1	24	1 LPER-STREF	P45440 streptomyce
20	21.5	18.7	19	1 ANP7-ELEGR	P11920 eleginus gr
21	21.5	18.7	20	1 MCRG-METTE	P22950 methanosarc
22	21	18.3	8	1 RS1 ERWCH	P37985 erwinitia chr
23	21	18.3	10	1 COXQ-THUOB	P80982 thunnus obe
24	21	18.3	12	1 H2AX-ONCMY	P83327 oncorhynch
25	21	18.3	13	1 PSPP-PINPS	P81668 pinus pinas
26	21	18.3	14	1 IF2G-RAT	P81795 rattus norv
27	21	18.3	16	1 FIBA-ODOHE	P14459 odocoileus
28	21	18.3	17	1 NEF HVLJ3	P12480 human immun
29	21	18.3	17	1 RM35-YEAST	P36530 saccharomyc
30	21	18.3	19	1 PHLC-STAIN	P80924 staphylococ
31	21	18.3	19	1 UP24-UPEIN	P82030 uperoleia i
32	21	18.3	20	1 CD4-SHEEP	P05542 oviss aries
33	21	18.3	20	1 COXQ-ONCMY	P80335 oncorhynch

```

34 20 17.4 9 1 DSIP RABIT
35 20 17.4 11 1 Q20A-COMTE
36 20 17.4 11 1 TKNA-SCYCA
37 20 17.4 13 1 PSAE-PEA
38 20 17.4 14 1 UC15 MAIZE
39 20 17.4 14 1 UHA1-CANFA
40 20 17.4 15 1 ATP2-SPIOL
41 20 17.4 16 1 ANP8-ELEGR
42 20 17.4 16 1 FIBA-FELCA
43 20 17.4 16 1 FIBA-HYLLA
44 20 17.4 19 1 FIBA-BISBO
45 20 17.4 19 1 PHSL-DESEN

```

## ALIGNMENTS

```

RESULT 1
MISG MISAN STANDARD; PRT; 21 AA.
AC P81474;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Misgurn.
OS Misgurnus anguillicaudatus (Oriental weatherfish) (Oriental
weatherloach).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cobitidae; Misgurnus.
OX NCBI_TaxID=75329;
RN [1]_SEQUENCE.
RP MEDLINE=97415401; PubMed=9271200;
RX Park C.B., Lee H.J., Park I.Y., Kim M.S., Kim S.C.;
RA "A novel antimicrobial peptide from the loach, Misgurnus
RT anguillicaudatus."
RL FEBS Lett. 411:173-178(1997).
CC -!- FUNCTION: Strong antimicrobial activity against several
CC Gram-positive and Gram-negative bacteria and fungi.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MASS SPECTROMETRY: MW=2502; METHOD=MALDI.
KW Antibiotic; Fungicide.
SQ SEQUENCE 21 AA; 2502 MW; 4A6E9D0AB391BCF1 CRC64;

Query Match 27.0%; Score 31; DB 1; Length 21;
Best Local Similarity 46.7%; Pred. No. 2.3e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 QFSKTAAGGAAAR 17
Db 6 ELSKFSKKGAAARR 20

RESULT 2
UC23 MAIZE STANDARD; PRT; 15 AA.
AC P80629;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Unknown protein from 2D-page of etiolated coleoptile (Spot 502)
DE (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]_SEQUENCE.
RP TISSUE=Coleoptile;
RC Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;

```

```

RT "The maize two dimensional gel protein database: towards an integrated
RL genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -1- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC protein is: 6.2, its MW is: 45.0 kDa.
DR Maize-2DPAGE; P80629; COLEOPTILE.
DR MaizeDB; 123955; -.
FT NON_TER 1 1
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1557 MW; C974ED33E9A4EC28 CRC64;

Query Match 23.5%; Score 27; DB 1; Length 15;
Best Local Similarity 50.0%; Pred. No. 6.3e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 14 AAERPGEA 21
| : : : :
Db 1 AGDKRPGDA 8

RESULT 3
YQK8_CABEL STANDARD; PRT; 22 AA.
AC Q03290;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 2.2 kDa protein C56G2.8 in chromosome III.
GN C56G2.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Connell M.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U23177; AAA64335.1; -.
DR PIR; TI5870; TI5870.
DR WormPep; C56G2.8; CE01877.
KW Hypothetical protein.
SQ SEQUENCE 22 AA; 2164 MW; 7585F9684FE19D51 CRC64;

Query Match 23.5%; Score 27; DB 1; Length 22;
Best Local Similarity 29.4%; Pred. No. 8.9e+02;
Matches 5; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 5 SKTAAKGFAAERPGEA 21
| : : : :
Db 4 AKSGGENVEGDADPGDA 20

RESULT 4
FIBA_LAMGL STANDARD; PRT; 18 AA.
AC F14454;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN FGA.
OS Lama glama (Llama), and
OS Lama vicugna (Vicugna) (Vicugna vicugna).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Lama.
OX NCBI_TaxID=9844, 9843;
RN [1]
RP SEQUENCE.
RC SPECIES=L.glama;
RA Blomback B., Blomback M., Grondahl N.J.;
RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
RN [2]
RP SEQUENCE.
RC SPECIES=L.vicugna;
RA Mross G.A., Doolittle R.F.;
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";
RL Arch. Biochem. Biophys. 122:674-684(1967).
CC -1- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -1- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC which cleaves fibrinopeptides A and B from alpha and beta chains,
CC and thus exposes the N-terminal polymerization sites responsible
CC for the formation of the soft clot.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 18 FIBRINOPEPTIDE A.
FT NON_TER 18 18
SQ SEQUENCE 18 AA; 1834 MW; 2444487B8B7F4CC6 CRC64;

Query Match 22.6%; Score 26; DB 1; Length 18;
Best Local Similarity 58.3%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 8 AAKGEAAERPG 19
| : : : :
Db 5 ADGGEFLAEGGG 16

RESULT 5
BOMN_BOMVA STANDARD; PRT; 24 AA.
ID BOMN_BOMVA
AC P01505;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Bombinin.
OS Bombina variegata (Yellow-bellied toad).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.
OX NCBI_TaxID=8348;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Csordas A., Michl H.;
RT "Isolation and structural resolution of a haemolytically active
RT polypeptide from the immune secretion of a European toad.";
RL Monatsh. Chem. 101:182-189(1970).
CC -1- FUNCTION: Has antimicrobial and hemolytic activities.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Skin.
CC -1- SIMILARITY: Belongs to the bombinin family.
DR PIR; A01766; BMTD.
KW Amphibian defense peptide; Antibiotic; Hemolysis; Amidation.
FT MOD_RES 24 24 AMIDATION.
SQ SEQUENCE 24 AA; 2294 MW; ACC0FCE3B7402E85 CRC64;

Query Match 21.7%; Score 25; DB 1; Length 24;
Best Local Similarity 53.8%; Pred. No. 1.9e+03;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAGGEA 13
| : : : :
Db 3 GALSAGALKGLA 15

```

```

CC      which cleaves fibrinopeptides A and B from alpha and beta chains,
CC      and thus exposes the N-terminal polymerization sites responsible
CC      for the formation of the soft clot.
KW      Blood coagulation; Plasma.
FT      PEPTIDE      1 16      FIBRINOPEPTIDE A.
FT      NON TER      16 16
SQ      SEQUENCE     16 AA; 1622 MW; 48598EB6292F4030 CRC64;

Query Match      20.9%; Score 24; DB 1; Length 16;
Best Local Similarity 50.0%; Pred. No. 1.8e+03;
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY      8 AAKGEAAERPG 19
DB      3 ATEGEFLAEGGG 14

RESULT 8
FIBA_PIG
ID      FIBA_PIG      STANDARD;      PRT;      17 AA.
AC      P14460;
DT      01-JAN-1990 (Rel. 13, Created)
DT      01-JAN-1990 (Rel. 13, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN      FGA.
OS      Sus scrofa (Pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX      NCBI_TaxID=9823;
RN      [1]
RP      SEQUENCE.
RA      Blomback B., Blomback M., Grondahl N.J.;
RT      "Studies on fibrinopeptides from mammals.";
RL      Acta Chem. Scand. 19:1789-1791(1965).
CC      -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC      polymerize into fibrin and acting as a cofactor in platelet
CC      aggregation.
CC      -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC      (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC      -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC      which cleaves fibrinopeptides A and B from alpha and beta chains,
CC      and thus exposes the N-terminal polymerization sites responsible
CC      for the formation of the soft clot.
DR      PIR; E29501; E29501.
KW      Blood coagulation; Plasma.
FT      PEPTIDE      1 16      FIBRINOPEPTIDE A.
FT      NON TER      17 17
SQ      SEQUENCE     17 AA; 1762 MW; 232EFEBB8B6B0A0C CRC64;

Query Match      20.9%; Score 24; DB 1; Length 17;
Best Local Similarity 60.0%; Pred. No. 1.9e+03;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      10 KGEAAERPG 19
DB      6 KGEFLAEGGG 15

RESULT 9
IBP2_PIG
ID      IBP2_PIG      STANDARD;      PRT;      23 AA.
AC      P24853;
DT      01-MAR-1992 (Rel. 21, Created)
DT      01-MAR-1992 (Rel. 21, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Insulin-like growth factor binding protein 2 (IGFBP-2) (IBP-2)
DE      (IGF-binding protein 2) (Fragment).
GN      IGFBP2.
OS      Sus scrofa (Pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX      NCBI_TaxID=9823;

```

```

RN  SEQUENCE.
RP  MEDLINE=92109718; PubMed=1722398;
RX  Coleman M.E., Pan Y.-C.E., Etherton T.D.;
RA  "Identification and NH2-terminal amino acid sequence of three
RT  insulin-like growth factor-binding proteins in porcine serum.";
RL  Biochem. Biophys. Res. Commun. 181:1131-1136(1991).
CC  -!- FUNCTION: IGF-binding proteins prolong the half-life of the IGFs
CC  and have been shown to either inhibit or stimulate the growth
CC  promoting effects of the IGFs on cell culture. They alter the
CC  interaction of IGFs with their cell surface receptors.
CC  -!- SUBCELLULAR LOCATION: Secreted.
CC  -!- MISCELLANEOUS: Binds IGF-II more than IGF-I.
CC  -!- SIMILARITY: Contains 1 IGFBP domain.
CC  -!- SIMILARITY: Contains 1 thyroglobulin type-I domain.
DR  PIR; JH0515; JH0515.
DR  InterPro; IPR000967; Insl_gro_fac_pr.
DR  InterPro; IPR000716; Thyroglobulin 1.
DR  PROSITE; PS00222; IGF BINDING; PARTIAL.
DR  PROSITE; PS00484; THYROGLOBULIN_1; PARTIAL.
KW  Growth factor binding.
FT  NON TER 23
SQ  SEQUENCE 23 AA; 2387 MW; 731104B88825D2C4 CRC64;

Query Match 20.9%; Score 24; DB 1; Length 23;
Best Local Similarity 66.7%; Pred. No. 2.5e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 14 AAERPGEAA 22
DB 12 AAXRPPPA 20
||| |||

RESULT 10
DBH3 RHILE STANDARD; PRT; 16 AA.
AC P80605;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA-binding protein H3-RL (Fragment).
OS Rhizobium leguminosarum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=384;
RN [1]
RP SEQUENCE.
RC STRAIN=CWR 538;
RX MEDLINE=20035754; PubMed=10570977;
RA Sojda J. III, Gu B., Lee J., Hoover T.R., Nixon B.T.;
RT "A rhizobial homolog of IHF stimulates transcription of dctA in
RT Rhizobium leguminosarum but not in Sinorhizobium meliloti.";
RL Gene 238:489-500(1999).
CC -!- FUNCTION: Histone-like DNA-binding protein which is capable of
CC wrapping DNA to stabilize it, and thus to prevent its denaturation
CC under extreme environmental conditions.
CC -!- SIMILARITY: Belongs to the bacterial histone-like protein family.
DR InterPro; IPR000119; Bac_DNAbind
DR Pfam; PF00216; Bac_DNA_binding 1.
DR PROSITE; PS00045; HISTONE LIKE; PARTIAL.
KW DNA-binding; DNA condensation.
FT NON TER 16
SQ SEQUENCE 16 AA; 1702 MW; F585ADED46B33964 CRC64;

Query Match 20.0%; Score 23; DB 1; Length 16;
Best Local Similarity 71.4%; Pred. No. 2.5e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 13 AAAERP 19
DB 9 AVAERAG 15
||| |||

```

```

RESULT 11
FIBA_BUBBU STANDARD; PRT; 19 AA.
AC P14442; P14443;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN FGA.
OS Bubalus bubalis (Domestic water buffalo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bubalus.
OX NCBI_TaxID=89462;
RN [1]
RP SEQUENCE.
RT Mross G.A., Doolittle R.F.;
RL "Amino acid sequence studies on artiodactyl fibrinopeptides.";
RN Arch. Biochem. Biophys. 122:674-684(1967).
[2]
RP SEQUENCE.
RC STRAIN=Italian breed;
RX MEDLINE=76040091; PubMed=1180969;
RA Balestrieri C., Colonna G., Itrace G.;
RT "Covalent structure of fibrinopeptides from buffaloes breeding in
RT Italy.";
RL Biochim. Biophys. Acta 405:517-521(1975).
CC -!- FUNCTION: Fibrinogen has a double function: Yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC which cleaves fibrinopeptides A and B from alpha and beta chains,
CC and thus exposes the N-terminal polymerization sites responsible
CC for the formation of the soft clot.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 19 FIBRINOPEPTIDE A.
FT VARIANT 8 8 S -> G (IN ITALIAN BRED).
FT NON TER 19
SQ SEQUENCE 19 AA; 1852 MW; 9BA41F0F55A54CC5 CRC64;

Query Match 20.0%; Score 23; DB 1; Length 19;
Best Local Similarity 50.0%; Pred. No. 2.9e+03;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 8 AAKCEAAARPDG 19
DB 6 AVSGEFLAEGG 17
||| |||

RESULT 12
COXO_THUOB STANDARD; PRT; 20 AA.
ID COXO_THUOB
AC P80983;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome c oxidase polypeptide VIII-1 (EC 1.9.3.1) (Fragment).
OS Thunnus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8241;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart, and Liver;
RX MEDLINE=97454291; PubMed=9310366;
RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and
RT liver.";

```

```

RL Eur. J. Biochem. 248:99-103(1997).
CC -!- FUNCTION: This protein is one of the nuclear-coded polypeptide
CC chains of cytochrome c oxidase, the terminal oxidase in
CC mitochondrial electron transport.
CC -!- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- SIMILARITY: Belongs to the cytochrome c oxidase VIII family.
DR PIR; S77991; S77991.
KW Oxidoreductase; Mitochondrion.
FT NON TER 20
SQ SEQUENCE 20 AA; 1993 MW; 5BFEA5080A862C58 CRC64;

Query Match 20.0%; Score 23; DB 1; Length 20;
Best Local Similarity 35.3%; Pred. No. 3e+03;
Matches 6; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 8 AAKGEAAARPCGAFAVA 24
DB 2 SAKPAKXXVTAGEQAIA 18

RESULT 13
P0RC_METTM STANDARD; PRT; 20 AA.
AC P80502;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Pyruvate synthase subunit P0RC (EC 1.2.7.1) (Pyruvate oxidoreductase
DE gamma chain) (POR) (Pyruvic-ferredoxin oxidoreductase gamma subunit)
DE (Fragment).
GN PORC.
OS Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=79929;
RN [1]
RP SEQUENCE.
RX MEDLINE=97261844; PubMed=9108258;
RA Tersteegen A., Linder D., Thauer R.K., Hedderich R.;
RT "Structures and functions of four anabolic 2-oxoacid oxidoreductases
RT in Methanobacterium thermoautotrophicum.";
RL Eur. J. Biochem. 244:862-868(1997).
CC -!- CATALYTIC ACTIVITY: Pyruvate + CoA + oxidized ferredoxin = acetyl-
CC CoA + CO(2) + reduced ferredoxin.
CC -!- COFACTOR: Binds 2 4Fe-4S clusters (By similarity).
CC -!- SUBUNIT: Heterotetramer of one alpha, one beta, one delta and one
CC gamma chain.
CC -!- MISCELLANEOUS: Has a pH optimum of 10.0 and an optimal temperature
CC of 80 degrees Celsius.
CC -!- SIMILARITY: Belongs to the bacterial-type ferredoxin family.
DR InterPro; IPR002869; POR.
KW Oxidoreductase; Electron transport; Iron-sulfur; Repeat; 4Fe-4S.
FT NON TER 20
SQ SEQUENCE 20 AA; 2167 MW; 481532134D42F353 CRC64;

Query Match 20.0%; Score 23; DB 1; Length 20;
Best Local Similarity 50.0%; Pred. No. 3e+03;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAQFSKTAAK 10
DB 10 GGQXAVTAAE 19

RESULT 14
FKB4_RAT STANDARD; PRT; 24 AA.
AC Q9QVC8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE FK506-binding protein 4 (EC 5.2.1.8) (Peptidyl-prolyl cis-trans

```

```

DE isomerase) (PPIase) (Rotamase) (p59 protein) (HSP binding
DE immunophilin) (HBI) (FKBP52 protein) (52 kDa FK506 binding protein)
DE (FKBP59) (Fragment).
GN FKBP4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC TISSUE=Leukemia;
RX MEDLINE=92285692; PubMed=1376003;
RA Tai P.-K.K., Albers M.W., Chang H., Faber L.E., Schreiber S.L.;
RT "Association of a 59-kilodalton immunophilin with the glucocorticoid
RT receptor complex.";
RL Science 256:1315-1318(1992).
CC -!- FUNCTION: Component of unactivated mammalian steroid receptor
CC complexes that sediment at 8-10 S. May have a rotamase activity.
CC May play a role in the intracellular trafficking of hetero-
CC oligomeric forms of steroid hormone receptors (By similarity).
CC -!- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline
CC (omega=0).
CC -!- SUBUNIT: Associates with HSP90 and HSP70 in unactivated steroid
CC hormone receptor complexes. Also interacts with peroxisomal
CC phytanoyl-CoA alpha-hydroxylase (PHYH). Interacts with NR3C1 and
CC dynein (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the FKBP-type PPIase family.
DR InterPro; IPR001179; FKBP_PPIase.
DR PROSITE; PS00453; FKBP_PPIASE_1; PARTIAL.
DR PROSITE; PS00454; FKBP_PPIASE_2; PARTIAL.
DR PROSITE; PS00059; FKBP_PPIASE_3; PARTIAL.
KW Isomerase; Rotamase; Nuclear protein.
FT UNSURE 10
FT UNSURE 14
FT UNSURE 18
FT UNSURE 20
FT NON TER 24
SQ SEQUENCE 24 AA; 2478 MW; 901A4B67B2AF5D16 CRC64;

Query Match 20.0%; Score 23; DB 1; Length 24;
Best Local Similarity 35.3%; Pred. No. 3.6e+03;
Matches 6; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 AQFSKTAAGGAARRP 18
DB 2 AEEMKVAENGAQSAPLP 18

RESULT 15
FKB4_CAMDR STANDARD; PRT; 18 AA.
AC F14444;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN FGA.
OS Camelus dromedarius (Dromedary) (Arabian camel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Tylopoda; Camelidae; Camelus.
OX NCBI_TaxID=9838;
RN [1]
RP SEQUENCE.
RX MEDLINE=67209145; PubMed=6033721;
RA Doolittle R.F., Schubert D., Schwartz S.A.;
RT "Amino acid sequence studies on artiodactyl fibrinopeptides. I.
RT Dromedary camel, mule deer, and cape buffalo.";
RL Arch. Biochem. Biophys. 118:456-467(1967).
CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS

```

CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.  
CC -1- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,  
CC which cleaves fibrinopeptides A and B from alpha and beta chains,  
CC and thus exposes the N-terminal polymerization sites responsible  
CC for the formation of the soft clot.  
KW Blood coagulation; Plasma.  
FT PEPTIDE 1 18 FIBRINOPEPTIDE A.  
FT NON TER 18 18  
SQ SEQUENCE 18 AA; 1835 MW; 244448763D7F4CC6 CRC64;  
  
Query Match 19.1%; Score 22; DB 1; Length 18;  
Best Local Similarity 50.0%; Pred. No. 3.8e+03;  
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
  
Qy 8 AAKGEAAERPG 19  
| : | | | |  
Db 5 ADEGEFLAEGGG 16  
  
Search completed: October 7, 2004, 17:54:02  
Job time : 17 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:37:47 ; Search time 72 Seconds  
(without alignments)

105.173 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAGGAARPGERAUA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 10229

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriap:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	25.2	15	10 Q9SQI6	Q9sqi6 oryza sativ
2	29	25.2	20	11 Q99JV2	Q99jv2 mus musculu
3	29	25.2	21	10 Q9S8K6	Q9s8k6 mesembryant
4	29	25.2	23	4 Q9BYM6	Q9bym6 homo sapien
5	28	24.3	17	3 Q00074	Q00074 aspergillus
6	28	24.3	20	2 Q51558	Q51558 pseudomonas
7	28	24.3	20	10 Q9SM31	Q9sm31 triticum ae
8	28	24.3	22	8 Q9T2I7	Q9t2i7 nicotiana s
9	27	23.5	21	4 Q86UF6	Q86uf6 homo sapien
10	27	23.5	22	11 Q8CJD5	Q8cjd5 rattus norv
11	27	23.5	22	12 Q85664	Q85664 reovirus (t
12	26	22.6	13	6 Q9TQ52	Q9tq52 equus cabal
13	26	22.6	15	5 Q9TW04	Q9tw04 crithidia f
14	26	22.6	18	11 Q55167	Q55167 rattus norv
15	26	22.6	20	2 Q85522	Q85522 chlamydia t
16	26	22.6	20	2 Q85523	Q85523 chlamydia t

17	26	22.6	20	2 Q85528	Q85528 chlamydia t
18	26	22.6	20	2 Q85512	Q85512 chlamydia t
19	26	22.6	20	2 Q85531	Q85531 chlamydia t
20	26	22.6	20	2 Q85510	Q85510 chlamydia t
21	26	22.6	20	2 Q85521	Q85521 chlamydia t
22	26	22.6	20	2 Q85514	Q85514 chlamydia t
23	26	22.6	20	2 Q85519	Q85519 chlamydia t
24	26	22.6	20	2 Q85517	Q85517 chlamydia t
25	26	22.6	20	2 Q85525	Q85525 chlamydia t
26	26	22.6	20	2 Q85526	Q85526 chlamydia t
27	26	22.6	20	2 Q69176	Q69176 chlamydia t
28	26	22.6	20	2 Q85520	Q85520 chlamydia t
29	26	22.6	20	2 Q85533	Q85533 chlamydia t
30	26	22.6	20	2 Q85509	Q85509 chlamydia t
31	26	22.6	20	2 Q85530	Q85530 chlamydia t
32	26	22.6	20	8 Q9T2I9	Q9t2i9 nicotiana s
33	26	22.6	20	15 Q78508	Q78508 human immun
34	26	22.6	21	6 Q9TR75	Q9tr75 bos taurus
35	26	22.6	22	5 Q94182	Q94182 caenorhabdi
36	26	22.6	22	6 Q9TRB7	Q9trb7 bos taurus
37	26	22.6	23	2 Q84P40	Q84f40 streptomyce
38	26	22.6	23	8 Q9T2S5	Q9t2s5 nicotiana s
39	25	21.7	8	5 P82688	P82688 periplaneta
40	25	21.7	13	2 Q9REI2	Q9rei2 acidiphiliu
41	25	21.7	16	3 Q9UR86	Q9ur86 candida par
42	25	21.7	19	6 Q9TR32	Q9tr32 ovis aries
43	25	21.7	19	13 Q91433	Q91433 gallus gall
44	25	21.7	20	8 Q9T2I8	Q9t2i8 nicotiana s
45	25	21.7	20	15 Q78505	Q78505 human immun

#### ALIGNMENTS

#### RESULT 1

Q9SQI6 PRELIMINARY; PRT; 15 AA.  
ID AC Q9SQI6;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Translation elongation factor (Fragment).  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Li Z.-Y., Chen S.-Y.;  
RT "Inducible expression of translation elongation factor 1A gene in rice seedlings in response to environmental stresses.";  
RL Acta Bot. Sin. 41:800-806(1999).  
DR EMBL; AF067195; AAC79991.1; -.  
DR Gramene; Q9SQI6; -.  
DR GO; GO:0003746; F:translation elongation factor activity; IEA.  
KW Elongation factor.  
FT NON TER 1  
SQ SEQUENCE 15 AA; 1514 MW; 95E5C46F069C9775 CRC64;

Query Match 25.2%; Score 29; DB 10; Length 15;  
Best Local Similarity 60.0%; Pred. No. 2e+03;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAK 10  
||: :|  
Db 4 GAKVTKAAAX 13

#### RESULT 2

Q99JV2 PRELIMINARY; PRT; 20 AA.  
ID AC Q99JV2  
Q99JV2

```

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005653; AAH05653.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 20 AA; 2223 MW; F203F6A80A7A0429 CRC64;

Query Match 25.2%; Score 29; DB 11; Length 20;
Best Local Similarity 62.5%; Pred. No. 2.7e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 13 AAAERPG 20
Db :|||:|
3 SAAKPGK 10

RESULT 3
Q9S8K6 PRELIMINARY; PRT; 21 AA.
ID Q9S8K6;
AC Q9S8K6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE 24 kDa salt-responsive glycoprotein/group 5 pathogenesis-related
DE protein homolog (Fragment).
OS Mesembryanthemum crystallinum (Common ice plant).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Alzooaceae; Mesembryanthemum.
OX NCBI_TaxID=3544;
RN [1]
RP SEQUENCE.
RX MEDLINE=95062715; PubMed=7972493;
RA Yen H.E., Edwards G.E., Grimes H.D.;
RT "Characterization of a salt-responsive 24-kilodalton glycoprotein in
RT Mesembryanthemum crystallinum."
RL Plant Physiol. 105:1179-1187(1994).
SQ SEQUENCE 21 AA; 2288 MW; B463C490647E076D CRC64;

Query Match 25.2%; Score 29; DB 10; Length 21;
Best Local Similarity 85.7%; Pred. No. 2.9e+03;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 AAAERPG 19
Db :|||:|
15 AAAARPG 21

RESULT 4
Q9BYM6 PRELIMINARY; PRT; 23 AA.
ID Q9BYM6;
AC Q9BYM6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE DSRNA adenosine deaminase (Fragment).
GN DSRNA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC TISSUE=Placenta;
RA Deblandre G., Marinx O., Nols C., Defrance P., Berr P., Huez G.,
RA Caput D.;
RT "The gene coding for the interferon-inducible human dsRNA adenosine
RT deaminase is transcribed into several messengers specifying different
RT proteins."
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; X98560; CAA67171.1; -.
FT NON_TER 23
SQ SEQUENCE 23 AA; 2543 MW; CF29179B7DFC1395 CRC64;

Query Match 25.2%; Score 29; DB 4; Length 23;
Best Local Similarity 50.0%; Pred. No. 3.1e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 10 KGEAAERPG 19
Db :|||:|
5 RNESAAGKPG 14

RESULT 5
Q00074 PRELIMINARY; PRT; 17 AA.
ID Q00074;
AC Q00074;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Extracellular acid protease (Fragment).
GN PEPA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=N400;
RX MEDLINE=95163121; PubMed=7532112;
RA Jarai G., Buxton F.;
RT "Nitrogen, carbon, and pH regulation of extracellular acidic proteases
RT of Aspergillus niger."
RL Curr. Genet. 26:238-244(1994).
DR EMBL; U03507; AAC48920.1; -.
DR MEROPS; A01.016; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
KW Protease.
FT NON_TER 17
SQ SEQUENCE 17 AA; 1694 MW; FF01D0EBC9FEAE0B CRC64;

Query Match 24.3%; Score 28; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 3.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 FSKTAA 9
Db :|||:|
4 FSKTAA 9

RESULT 6
Q51558 PRELIMINARY; PRT; 20 AA.
ID Q51558;
AC Q51558;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Protein 10 (77).
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90130290; PubMed=2105298;
RA Hamood A.N., Iglewski B.H.;

```



RT "Expression of the Pseudomonas aeruginosa toxA positive regulatory  
 RT gene (regA) in Escherichia coli.";  
 RL 7. Bacterial. 172:589-594 (1990).  
 DR EMBL; M33819; AAA25982.1; -.  
 SQ SEQUENCE 20 AA; 2158 MW; 8878365B519BFE4A CRC64;

Query Match 24.3%; Score 28; DB 2; Length 20;  
 Best Local Similarity 37.5%; Pred. No. 3.8e+03;  
 Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 5 SNTAAKGEAAARPGGE 20  
 : : : : :  
 Db 2 ARIRAGSSRACRLGE 17

## RESULT 7

Q9SM31 PRELIMINARY; PRT; 20 AA.  
 AC Q9SM31;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE Putative sucrose synthase (Fragment).  
 GN SS1.  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;  
 OC Triticeae; Triticum.  
 OX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Chinese spring;  
 RA Helguera M., Khan I.A., Dubcovsky J.;  
 RT "Development of PCR markers for wheat leaf rust resistance gene  
 Lr47.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ249623; CAB61261.1; -.  
 FT NON TER 1  
 SQ SEQUENCE 20 AA; 2298 MW; CC4EFE0F3877CCDC CRC64;

Query Match 24.3%; Score 28; DB 10; Length 20;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+03;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 ERPGEA 21  
 : : : : :  
 Db 15 EQPGEA 20

## RESULT 8

Q9T217 PRELIMINARY; PRT; 22 AA.  
 AC Q9T217;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE 14.3 photosystem I PSAE protein (Fragment).  
 OS Nicotiana glauca (Wood tobacco).  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Lamiales; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4096;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=94105345; PubMed=8278548;  
 RA Obokata J., Mikami K., Hayaashida N., Nakamura M., Sugiura M.;  
 RT "Molecular heterogeneity of photosystem I. psaD, psaE, psaF, psaH, and  
 psaL are all present in isoforms in Nicotiana spp.";  
 RL Plant Physiol. 102:1259-1267 (1993).  
 SQ SEQUENCE 22 AA; 2033 MW; 488E17845223FFF7 CRC64;

Query Match 24.3%; Score 28; DB 8; Length 22;

Best Local Similarity 38.9%; Pred. No. 4.2e+03;  
 Matches 7; Conservative 1; Mismatches 10; Indels 0; Gaps 0;  
 QY 6 KTAAGGEAAARPGGAAY 23  
 : : : : :  
 Db 3 EAAPPAATAAPAPV 20

## RESULT 9

Q86UF6 PRELIMINARY; PRT; 21 AA.  
 AC Q86UF6;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Betacellulin (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nakagawa T.;  
 RT "The 5' end of human betacellulin gene and molecular scanning for  
 RT mutations in Japanese patients with type2 diabetes mellitus.";  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB109442; BAC76610.1; -.  
 FT NON TER 21  
 SQ SEQUENCE 21 AA; 2083 MW; 8387517EAA8B9FEB CRC64;

Query Match 23.5%; Score 27; DB 4; Length 21;  
 Best Local Similarity 41.2%; Pred. No. 5.5e+03;  
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 8 AAKGEAAARPGGAAY 24  
 : : : : :  
 Db 4 AARGGASSLPALLALA 20

## RESULT 10

Q8CJD5 PRELIMINARY; PRT; 22 AA.  
 AC Q8CJD5;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Podocin (Fragment).  
 GN NPHS2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Morita H., Yoshimura A., Makino R., Inui K., Nakao N., Usami T.,  
 RA Roselli S., Antignac C., Matsuyama M., Ideuka T.;  
 RT "Rat genome fragment containing a part of exons and all of the 3'UTR  
 of Nphs2 as well as microsatellite sites.";  
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB094123; BAC23093.1; -.  
 FT NON TER 1  
 FT NON TER 22  
 SQ SEQUENCE 22 AA; 2313 MW; DBC687F825956D93 CRC64;

Query Match 23.5%; Score 27; DB 11; Length 22;  
 Best Local Similarity 75.0%; Pred. No. 5.8e+03;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 8 AAKGEAA 15  
 : : : : :  
 Db 11 AAEGEKAA 18

```

RESULT 11
Q85664
ID Q85664 PRELIMINARY; PRT; 22 AA.
AC Q85664;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Major core protein lambda-1 (Fragment).
OS Reovirus (type 3 / strain Dearing).
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus.
OX NCBI_TaxID=10886;
RN [1]
RP SEQUENCE OF 1-6 FROM N.A.
RC STRAIN=Dearing;
RX MEDLINE=82217029; PubMed=7086967;
RA Kozak M.;
RT "Sequences of ribosome binding sites from the large size class of reovirus mRNA.";
RL J. Virol. 42:467-473 (1982).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Dearing;
RX MEDLINE=83017876; PubMed=6927854;
RA Antczak J.B., Chmelo R., Pickup D.J., Joklik W.K.;
RT "Sequences at both termini of the 10 genes of reovirus serotype 3 (strain Dearing).";
RL Virology 121:307-319 (1982).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Dearing;
RX MEDLINE=83017877; PubMed=7123853;
RA Gaillard R.K., Li J.K., Keene J.D., Joklik W.K.;
RT "The sequences at the termini of four genes of the three reovirus serotypes.";
RL Virology 121:320-326 (1982).
DR EMBL; J02317; AAA47254.1; -.
FT NON_TER 22
SQ SEQUENCE 22 AA; 2462 MW; D6706A861D930E1F CRC64;

Query Match 23.5%; Score 27; DB 12; Length 22;
Best Local Similarity 41.7%; Pred. No. 5.8e+03;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 KTAAGGEAAER 17
Db 11 KSSGKNDSTER 22

RESULT 12
Q9TQS2
ID Q9TQS2 PRELIMINARY; PRT; 13 AA.
AC Q9TQS2;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE Transferrin (Fragment).
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Giffard J.M., Brandon R.B., Bell T.K.;
RT "Further identification of single nucleotide polymorphisms in the equine transferrin gene.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF185752; AAF05466.1; -.
DR EMBL; AF185748; AAF05462.1; -.
DR EMBL; AF185749; AAF05463.1; -.
DR EMBL; AF185750; AAF05464.1; -.
DR EMBL; AF185751; AAF05465.1; -.
FT NON_TER 1
FT NON_TER 13

RESULT 13
Q9TWU4
ID Q9TWU4 PRELIMINARY; PRT; 15 AA.
AC Q9TWU4;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE KDNA-binding protein (Fragment).
OS Crithidia fasciculata.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Crithidia.
OX NCBI_TaxID=5656;
RN [1]
RP SEQUENCE.
RX MEDLINE=93245912; PubMed=8482355;
RA Tittawella I.;
RT "Identification of DNA-binding proteins in the parasitic protozoan Crithidia fasciculata and evidence for their association with the mitochondrial genome.";
RL Exp. Cell Res. 206:143-151 (1993).
DR PIR; A49177; A49177.
SQ SEQUENCE 15 AA; 1453 MW; 7C737E529D9D8270 CRC64;

Query Match 22.6%; Score 26; DB 5; Length 15;
Best Local Similarity 46.2%; Pred. No. 5.5e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 12 EAAAEPPGEAAVA 24
Db 1 DAPASAPKAAA 13

RESULT 14
Q55167
ID Q55167 PRELIMINARY; PRT; 18 AA.
AC Q55167;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE ARE1 (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEW.1W/Gun.
RX MEDLINE=99009314; PubMed=9790748;
RA Walter L., Guenther E.;
RT "Identification of a novel highly conserved gene in the centromeric part of the major histocompatibility complex.";
RL Genomics 52:298-304 (1998).
DR EMBL; AJ223831; CAA11568.1; -.
FT NON_TER 18
SQ SEQUENCE 18 AA; 1505 MW; B88156A2048C3388 CRC64;

Query Match 22.6%; Score 26; DB 11; Length 18;
Best Local Similarity 53.3%; Pred. No. 6.6e+03;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 8 AAKGEAAAEPPGEAA 22
Db 1 DAPASAPKAAA 13

```

Db 4 AATMAAAAGAGVAA 18

## RESULT 15

O85522 PRELIMINARY; PRT; 20 AA.  
AC O85522;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)  
DE Major outer membrane protein (Fragment).  
GN omp-1.  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=813;  
RN [1]\_  
RP SEQUENCE FROM N.A.  
RC STRAIN=B23a;  
RA Hsieh Y.-H., Bobo L.D.;  
RT "Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania.";  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF070275; AAC25245.1; -.  
FT NON\_TER 1  
FT NON\_TER 20  
SQ SEQUENCE 20 AA; 1929 MW; 504C8985119AF1F3 CRC64;

Query Match 22.6%; Score 26; DB 2; Length 20;  
Best Local Similarity 66.7%; Pred. No. 7.3e+03;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 7 TAAKGEAAA 15  
|||  
Db 4 TATTGNAAA 12

Search completed: October 7, 2004, 17:56:33  
Job time : 76 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:36:12 ; Search time 84.5 Seconds  
(without alignments)  
80.250 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTRAKGEAAERPGEEAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 538975

Minimum DB seq length: 0  
Maximum DB seq length: 24

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1980s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	24	3	AAY95896 MANS pept
2	115	100.0	24	6	ABP97765 Myristoyl
3	39	33.9	24	4	ABB40149 Peptide #
4	39	33.9	24	4	AAM33792 Peptide #
5	39	33.9	24	4	AAM73597 Human bon
6	39	33.9	24	4	AAM60912 Human bon
7	39	33.9	24	4	ABG55330 Human liv
8	39	33.9	24	5	ABG43467 Human pep
9	37	32.2	16	2	AAY30222 KEA16, am
10	37	32.2	16	2	AAY32964 Membrane
11	37	32.2	16	7	AAE38860 Membrane
12	37	32.2	19	2	AAY20179 Human bet
13	37	32.2	24	7	ADB81153 Human pan
14	36	31.3	20	2	AAY36444 Fragment
15	36	31.3	20	6	ADAL1979 Human nov
16	35	30.4	21	6	ABU06998 Maize Sta
17	35	30.4	21	6	ABU06996 Maize Sta
18	34	29.6	16	2	AAR04447 Human imm
19	34	29.6	16	2	AAW30220 KEA16, am
20	34	29.6	16	2	AAV32962 Membrane
21	34	29.6	16	7	AAE38858 Membrane
22	34	29.6	21	2	AAW19062 Trypanoso
23	34	29.6	21	2	AAW19087 Trypanoso
24	34	29.6	21	2	AAV32839 TCE repea
25	34	29.6	21	2	AAV23311 Epitope o

## ALIGNMENTS

## RESULT 1

AY95896	26	34	29.6	21	2	AY23313 Repeat se
ID AAY95896 standard; peptide; 24 AA.	27	34	29.6	21	3	AB26466 T.cruzi T
XX	28	34	29.6	21	3	AB26468 T.cruzi s
AC AAY95896;	29	34	29.6	23	4	AU03178 Partial p
XX	30	33	28.7	15	3	AY58971 Copeptide
DT 20-NOV-2000 (first entry)	31	33	28.7	15	3	AY82060 MHC class
XX	32	33	28.7	15	4	AG63202 Peptide w
DE MANS peptide inhibitor of MARCKS-related mucus secretion.	33	33	28.7	16	2	AAW30841 MyokL pro
XX	34	33	28.7	16	2	AAW30221 EXA16, am
MANS; MARCKS; myristoylated alanine-rich C kinase substrate; human;	35	33	28.7	16	2	AAW30223 KEA16, am
KW mucus secretion; inhibitor; bronchitis; asthma; cystic fibrosis;	36	33	28.7	16	2	AY32963 Membrane
KW chronic obstructive pulmonary disease; pneumonia; empysema; influenza;	37	33	28.7	16	2	AY32965 Membrane
XX	38	33	28.7	16	5	ABG92842 Synthetic
OS Homo sapiens.	39	33	28.7	16	5	ABG92827 Synthetic
XX	40	33	28.7	16	5	ABG92508 Self-asse
PN WO200050062-A2.	41	33	28.7	16	5	ABG92493 Self-asse
XX	42	33	28.7	16	6	ABP59360 Self-asse
PD 31-AUG-2000.	43	33	28.7	16	6	ABP59376 Self-asse
XX	44	33	28.7	16	7	AAE38859 Membrane
PF 24-FEB-2000; 2000WO-US005050.	45	33	28.7	16	7	AAE38861 Membrane
XX						
PR 24-FEB-1999; 99US-00256154.						
XX						
PA (UYNC-) UNIV NORTH CAROLINA STATE.						
XX						
PI Li Y, Martin LD, Adler KB;						
XX						
DR WPI; 2000-572036/53.						
XX						
PT Regulating mucus secretion by a mucus-secreting cell, useful for treating						
PT e.g. bronchitis, asthma or pneumonia, by administering a compound that						
PT inhibits or enhances myristoylated alanine-rich C-kinase substrate						
XX						
PS protein.						
XX						
PS Claim 12; Page 40; 66pp; English.						
XX						
CC The present sequence is that of MANS peptide, comprising the N-terminal						
CC region of human myristoylated alanine-rich C kinase substrate MARCKS						
CC protein (see AAY95896), a major cellular substrate for protein kinase S.						
CC MANS peptide inhibits secretion of mucus from mucus membranes and mucus-						
CC secreting cells, including human airway epithelial cells. It is suggested						
CC to block attachment of MARCKS protein to the mucin granule, thus blocking						
CC or inhibiting the release of mucin granules and the secretion of mucus by						
CC the cell. The invention relates to methods and compounds for decreasing						
CC mucus secretion, particularly in the airways. Such compounds include MANS						
CC peptide and antisense oligonucleotides to MARCKS. They are useful in						

CC inhibiting mucus secretion in conditions such as bronchitis, cystic  
CC fibrosis, chronic obstructive pulmonary disease, asthma, emphysema,  
CC pneumonia, influenza, rhinitis and the common cold

XX SQ Sequence 24 AA;

Query Match 100.0%; Score 115; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 9.6e-11;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAERPGAAVA 24  
|||  
DB 1 GAQFSKTAAGGAAERPGAAVA 24

RESULT 2

ABP97765  
ID ABP97765 standard; peptide; 24 AA.

XX AC

ABP97765;

XX DT 11-AUG-2003 (first entry)

XX DE Myristoylated N-terminal sequence (MANS) peptide.

XX KW Myristoylated N-terminal sequence; MANS; MARCKS; mucus secretion;  
KW inflammatory mediator; inflammation; respiratory disease; asthma;  
KW chronic bronchitis; chronic obstructive pulmonary disease; COPD;  
KW bowel disease; irritable bowel syndrome; Crohn's disease;  
KW ulcerative colitis; skin disease; rosacea; eczema; psoriasis; acne;  
KW autoimmune disease; pain; arthritis; cystic fibrosis.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal myristate chain attached"

XX WO2003000027-A2.

XX PD 03-JAN-2003.

XX PF 26-JUN-2002; 2002WO-US022270.

XX PR 26-JUN-2001; 2001US-0300933P.

XX PA (UYN-C) UNIV NORTH CAROLINA STATE.

XX PI Martin LD, Adler KB, Li Y;

XX DR WPI; 2003-278239/27.

XX PT Method of regulating inflammation comprises administering a composition  
XX comprising a MANS peptide or an active fragment thereof.

XX PS Disclosure; Page 26; 54pp; English.

XX CC The present sequence represents a myristoylated N-terminal sequence  
XX (MANS) peptide. The MANS peptide is identical to the first 24 amino acids  
XX of MARCKS, and mediates insertion of MARCKS into membranes. The MANS  
XX peptide inhibits both mucus secretion and inflammatory mediators. The  
XX specification describes a method of regulating inflammation. The method  
XX comprises administering a composition comprising a MANS peptide. The  
XX method is useful for treating inflammation caused by respiratory diseases  
XX (e.g. asthma, chronic bronchitis and chronic obstructive pulmonary  
XX disease (COPD), bowel diseases (e.g. irritable bowel syndrome, Crohn's  
XX disease and ulcerative colitis), skin diseases (e.g. rosacea, eczema,  
XX psoriasis and severe acne), autoimmune diseases and pain syndromes,  
XX arthritis and cystic fibrosis

XX SQ Sequence 24 AA;

Query Match 100.0%; Score 115; DB 6; Length 24;

Best Local Similarity 100.0%; Pred. No. 9.6e-11;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAQFSKTAAGGAAERPGAAVA 24  
|||  
DB 1 GAQFSKTAAGGAAERPGAAVA 24

RESULT 3

ABB40149

ID ABB40149 standard; peptide; 24 AA.

XX AC

ABB40149;

XX DT 04-FEB-2002 (first entry)

XX DE Peptide #7655 encoded by human foetal liver single exon probe.

XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PR 30-JAN-2001; 2001WO-US000669.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483447/52.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing  
XX gene expression in human fetal liver.

XX PS Claim 27; SEQ ID NO 32784; 639pp + Sequence Listing; English.

XX CC The invention relates to a single exon nucleic acid probe for measuring  
XX human gene expression in a sample derived from human foetal liver. The  
XX single exon nucleic acid probes may be used for predicting, measuring and  
XX displaying gene expression in samples derived from human fetal liver. The  
XX present sequence is a peptide encoded by a single exon nucleic acid probe  
XX of the invention. Note: The sequence data for this patent did not form  
XX part of the printed specification, but was obtained in electronic format  
XX directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 24 AA;

Query Match 33.9%; Score 39; DB 4; Length 24;

Best Local Similarity 50.0%; Pred. No. 67;

Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 4 FSKTAAGGAAERPG 19

DB 4 FSHVFNSTARGKEPG 19

RESULT 4

AAM33792

ID AAM33792 standard; protein; 24 AA.

XX AC

AAM33792;

DT 17-OCT-2001 (first entry)  
 XX Peptide #7829 encoded by probe for measuring placental gene expression.  
 DE  
 XX  
 XX Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157272-A2.  
 XX  
 XX 09-AUG-2001.  
 PD  
 XX  
 XX 30-JAN-2001; 2001WO-US000663.  
 XX  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI  
 XX WPI; 2001-488897/53.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human placenta.  
 PT  
 XX Claim 27; SEQ ID NO 34061; 654pp; English.  
 PS  
 XX The present invention relates to single exon nucleic acid probes (SENP:  
 CC see AA1315-AA15746). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders  
 CC  
 XX Sequence 24 AA;  
 SQ  
 Query Match 33.9%; Score 39; DB 4; Length 24;  
 Best Local Similarity 50.0%; Pred. No. 67;  
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 FSKTAAGGAAERPG 19  
 DB || | | | | | |  
 4 FSHVPANSTAAGEKPG 19  
 RESULT 5  
 AAM73597  
 ID AAM73597 standard; protein; 24 AA.  
 XX  
 AC AAM73597;  
 XX  
 XX 06-NOV-2001 (first entry)  
 DT  
 XX Human bone marrow expressed probe encoded protein SEQ ID NO: 33903.  
 DE  
 XX Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157276-A2.  
 XX  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US000668.  
 PF  
 XX

PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI  
 XX WPI; 2001-488900/53.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human bone marrow.  
 PT  
 XX Example 4; SEQ ID NO 33903; 658pp + Sequence Listing; English.  
 PS  
 XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention  
 CC  
 XX Sequence 24 AA;  
 SQ  
 Query Match 33.9%; Score 39; DB 4; Length 24;  
 Best Local Similarity 50.0%; Pred. No. 67;  
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 FSKTAAGGAAERPG 19  
 DB || | | | | | |  
 4 FSHVPANSTAAGEKPG 19  
 RESULT 6  
 AAM60912  
 ID AAM60912 standard; protein; 24 AA.  
 XX  
 AC AAM60912;  
 XX  
 XX 05-NOV-2001 (first entry)  
 DT  
 XX Human brain expressed single exon probe encoded protein SEQ ID NO: 33017.  
 DE  
 XX Human; brain expressed exon; gene expression analysis; probe; microarray;  
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US000667.  
 PF  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI  
 XX WPI; 2001-483446/52.  
 DR  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 PT

PT brains.  
 PS Example 4; SEQ ID NO 33017; 650pp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is a protein encoded by one of  
 CC the probes of the invention  
 XX  
 SQ Sequence 24 AA;  
 Query Match 33.9%; Score 39; DB 4; Length 24;  
 Best Local Similarity 50.0%; Pred. No. 67;  
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 FSKTAAKGAAERPG 19  
 || | | | | | | | |  
 DB 4 FSHVPANSTAGEKPG 19  
 || | | | | | | | |  
 RESULT 7  
 ABG55330  
 ID ABG55330 standard; peptide; 24 AA.  
 AC ABG55330;  
 XX  
 DT 25-FEB-2003 (first entry)  
 XX  
 DE Human liver peptide, SEQ ID NO 33978.  
 KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;  
 KW hypercholesterolaemia; coronary heart disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157273-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US0000664.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488898/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human adult liver.  
 XX  
 PS Claim 27; SEQ ID NO 33978; 658pp; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for  
 CC measuring human gene expression in a sample derived from human adult  
 CC liver, comprising one of 13109 defined nucleotide sequences given in the  
 CC specification (or complements/ fragments). The probe hybridises at high  
 CC stringency to a nucleic acid molecule expressed in the human adult liver.  
 CC (I) may be used for predicting, measuring and displaying gene expression  
 CC in samples derived from human adult liver. The genes identified may be  
 CC involved in genetic liver diseases such as cirrhosis,  
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
 CC associated with coronary heart disease. ABG47348-ABG59930 represent human

CC liver single exon encoded peptides of the invention. Note: The sequence  
 CC information for this patent does not appear in the printed specification  
 CC but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 24 AA;  
 Query Match 33.9%; Score 39; DB 4; Length 24;  
 Best Local Similarity 50.0%; Pred. No. 67;  
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 FSKTAAKGAAERPG 19  
 || | | | | | | | |  
 DB 4 FSHVPANSTAGEKPG 19  
 || | | | | | | | |  
 RESULT 8  
 ABG43467  
 ID ABG43467 standard; peptide; 24 AA.  
 AC ABG43467;  
 XX  
 DT 19-AUG-2002 (first entry)  
 XX  
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 33132.  
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200186003-A2.  
 XX  
 PD 15-NOV-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US0000665.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2002-114183/15.  
 XX  
 PT Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples.  
 XX  
 PS Claim 27; SEQ ID NO 33132; 634pp; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of probes  
 CC ; the novel set of probes which hybridise at high stringency to a nucleic  
 CC acid expressed in the human lung; measuring gene expression in a sample  
 CC derived from human lung, comprising (a) contacting the array with a  
 CC collection of detectably labeled nucleic acids derived from human lung





XX Cell cultures utilizing stable macroscopic membranes formed by the self-  
 PT assembly of amphiphilic peptides.  
 XX  
 XX Example 5; Col 26; 49pp; English.  
 XX  
 CC This sequence represents a membrane forming amphiphilic peptide used in  
 CC the in vitro method of the invention. The method is for culturing cells  
 CC utilising stable macroscopic membranes formed by the self-assembly of  
 CC amphiphilic peptides. The peptides are salt-induced to form insoluble and  
 CC protease-resistant protein filaments with a beta-sheet secondary  
 CC structure. The membranes may be useful in a wide variety of  
 CC medical, research, industrial and biomaterial applications such as slow-  
 CC diffusion drug delivery systems, artificial skin and separation matrices.  
 CC The membranes may be used to support in vitro cell attachment and growth  
 CC and for supporting artificial tissue (e.g. for in vivo use as implants).  
 CC They are particularly useful as experimental models for Alzheimer's  
 CC disease and scrapie infection and so may be used in disease modelling  
 CC experiments and to assay for agents which modulate the disease processes.  
 CC Additionally, they may be used in this way to study liver cirrhosis,  
 CC kidney amyloidosis and other protein conformational diseases. The  
 CC membrane in the membrane/cell mixture produced in the method: (i)  
 CC supports cellular attachment and growth; (ii) self-assembles to form  
 CC large, macroscopic membranes that are insoluble and stable in aqueous  
 CC solutions, serum and ethanol; (iii) is highly resistant to heat,  
 CC alkaline/acidic pHs, chemical denaturants and proteolytic digestion; (iv)  
 CC is non-cytotoxic and non-immunogenic; (v) is visible to the naked eye if  
 CC dyed but is otherwise transparent; (vi) may form thin, permeable, high  
 CC density sheets or fibril like structures with simple structures, high  
 CC tensile strength and a porous structure; (vii) may be metabolised by  
 CC humans and animals; (viii) is inexpensive to produce; and (ix) can be  
 CC produced and stored in sterile conditions  
 XX  
 SQ Sequence 16 AA;

Query Match 32.2%; Score 37; DB 2; Length 16;  
 Best Local Similarity 56.2%; Pred. No. 89;  
 Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 KTAAGGEAAERPGEA 21  
 DB 1 KAEAKAEAKAEAKAEA 16

RESULT 11  
 AAEE38860  
 ID AAE38860 standard; peptide; 16 AA.  
 XX  
 AC AAE38860;

DE 18-DEC-2003 (first entry)  
 XX Membrane forming peptide, KEA16.  
 XX  
 KW Macroscopic membrane; biomaterial; medical product; internal lining;  
 KW artificial skin; slow-diffusion drug delivery; chromatography column;  
 KW separation matrix; suture; dialysis membrane; viral filter; fabric;  
 KW pathology; Alzheimer's disease; scrapie infection.  
 XX  
 OS Unidentified.  
 XX  
 PN US6548630-B1.  
 XX  
 PD 15-APR-2003.  
 XX  
 PF 22-JUL-1997; 97US-00898300.  
 XX  
 PR 28-DEC-1992; 92US-00973326.  
 PR 30-NOV-1994; 94US-00346849.  
 XX  
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.  
 XX  
 PI Zhang S, Lockshin C, Rich A, Holmes T;

XX WPI; 2003-584339/55.  
 DR  
 XX  
 PT Macroscopic membrane useful in biomaterial applications e.g. sutures is  
 PT formed by self-assembly of amphiphilic peptides in an aqueous solution  
 PT containing monovalent metal cations.  
 XX  
 XX Example 5; Col 25; 51pp; English.  
 XX  
 CC The invention relates to a macroscopic membrane formed by self-assembly  
 CC of amphiphilic peptides in an aqueous solution containing monovalent  
 CC metal cations. The peptides have alternating hydrophobic and hydrophilic  
 CC amino acids and are complementary and structurally compatible. The  
 CC membranes are useful as biomaterial for medical products (e.g. sutures,  
 CC artificial skin, internal linings), as vehicles for slow-diffusion drug  
 CC delivery (preferably for protein type drugs e.g. erythropoietin, tissue  
 CC type plasminogen activator, synthetic haemoglobin and insulin), as  
 CC separation matrices (e.g. dialysis membranes, chromatography columns),  
 CC filters for the removal of viruses and other microscopic contaminants),  
 CC for other uses requiring permeable and water-insoluble material (e.g. for  
 CC culturing cell monolayers including differentiated cells and/or  
 CC stratified cell layers), for the preparation of very thin and transparent  
 CC fabric and as a model system for investigating the properties of  
 CC biological protein structures and providing insights into the pathology  
 CC and potential treatment of conditions involving the presence of the  
 CC proteins and proteinaceous structures (e.g. Alzheimer's disease, scrapie  
 CC infection and in origin of life studies related to cell membranes and  
 CC cellular compartmentalisation). The present sequence is a membrane  
 CC forming peptide used in the invention  
 XX  
 SQ Sequence 16 AA;

Query Match 32.2%; Score 37; DB 7; Length 16;  
 Best Local Similarity 56.2%; Pred. No. 89;

Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 KTAAGGEAAERPGEA 21  
 DB 1 KAEAKAEAKAEAKAEA 16

RESULT 12  
 AAY20179  
 ID AAY20179 standard; protein; 19 AA.  
 XX  
 AC AAY20179;

XX 22-JUL-1999 (first entry)

DE Human beta-amyloid precursor protein mutant fragment 4.

XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;  
 KW frameshift mutation; age-related disease; neurodegenerative disorder;  
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;  
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;  
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;  
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;  
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;  
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;  
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;  
 KW high mobility group protein-C; neuroendocrine specific protein A.

XX Synthetic.  
 OS Homo sapiens.  
 XX WO9845322-A2.  
 XX  
 PN 15-OCT-1998.  
 XX  
 PD 02-APR-1998; 98WO-IB000705.  
 XX  
 PF 10-APR-1997; 97US-0043163P.  
 XX  
 PR





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:54:08 ; Search time 83.5 Seconds  
(without alignments)  
92.493 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAKGEAAERPGAAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 284360

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	100.0	24	12	US-10-180-753-1
2	115	100.0	24	16	US-10-802-644-1
3	39	33.9	24	9	US-09-864-761-44320
4	37	32.2	16	16	US-10-390-472-35
5	37	32.2	20	10	US-09-939-537-52
6	36	31.3	20	12	US-10-372-876-507
7	36	31.3	20	14	US-10-097-065-507
8	35	30.4	21	16	US-10-109-048-902
9	35	30.4	21	16	US-10-109-048-904
10	34	29.6	16	16	US-10-390-472-33
11	33	28.7	15	9	US-09-765-301-28
12	33	28.7	15	10	US-09-765-644A-28
13	33	28.7	15	15	US-10-438-538-32
14	33	28.7	16	9	US-09-778-200-6
15	33	28.7	16	9	US-09-778-200-21

16	33	28.7	16	14	US-10-192-832-6	Sequence 6, Appli
17	33	28.7	16	14	US-10-192-832-22	Sequence 22, Appl
18	33	28.7	16	16	US-10-390-472-34	Sequence 34, Appl
19	33	28.7	16	16	US-10-390-472-36	Sequence 36, Appl
20	32	27.8	16	9	US-09-778-200-2	Sequence 2, Appli
21	32	27.8	16	14	US-10-192-832-2	Sequence 2, Appli
22	32	27.8	20	12	US-10-283-017-2058	Sequence 2058, Ap
23	32	27.8	21	10	US-09-974-879-413	Sequence 413, App
24	32	27.8	21	10	US-09-305-736-413	Sequence 413, App
25	32	27.8	21	11	US-09-818-683-413	Sequence 413, App
26	32	27.8	21	12	US-10-621-401-413	Sequence 413, App
27	32	27.8	21	12	US-09-818-683-413	Sequence 413, App
28	32	27.8	21	15	US-10-448-163-28	Sequence 28, Appl
29	32	27.8	22	12	US-10-619-778-7	Sequence 7, Appli
30	32	27.8	22	12	US-09-726-792A-7	Sequence 7, Appli
31	32	27.8	22	15	US-10-462-138-6	Sequence 6, Appli
32	32	27.8	22	16	US-10-755-784-7	Sequence 7, Appli
33	31.5	27.4	20	10	US-09-895-298-150	Sequence 150, App
34	31	27.0	15	9	US-09-765-301-10	Sequence 10, Appl
35	31	27.0	15	9	US-09-765-301-24	Sequence 24, Appl
36	31	27.0	15	10	US-09-765-644A-10	Sequence 10, Appl
37	31	27.0	15	10	US-09-765-644A-24	Sequence 24, Appl
38	31	27.0	15	15	US-10-438-538-14	Sequence 14, Appl
39	31	27.0	15	15	US-10-438-538-28	Sequence 28, Appl
40	31	27.0	16	10	US-09-847-586-6	Sequence 6, Appli
41	31	27.0	16	10	US-09-847-586-14	Sequence 14, Appl
42	31	27.0	16	10	US-09-847-586-15	Sequence 15, Appl
43	31	27.0	16	10	US-09-847-586-16	Sequence 16, Appl
44	31	27.0	16	10	US-09-847-586-17	Sequence 17, Appl
45	31	27.0	16	10	US-09-847-586-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1  
US-10-180-753-1  
; Sequence 1, Application US/10180753  
; Publication No. US20030013652A1  
; GENERAL INFORMATION:  
; APPLICANT: Martin, Linda  
; APPLICANT: Adler, Kenneth  
; APPLICANT: Li, Yuehua  
; TITLE OF INVENTION: BLOCKING PEPTIDE FOR INFLAMMATORY CELL SECRETION  
; FILE REFERENCE: 5051.574  
; CURRENT APPLICATION NUMBER: US/10/180,753  
; CURRENT FILING DATE: 2002-06-26  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 1  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic peptide  
; FEATURE:  
; NAME/KEY: MOD RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: N-terminal myristoylation  
US-10-180-753-1

Query Match 100.0%; Score 115; DB 12; Length 24;  
Best Local Similarity 100.0%; Pred. No. 8.6e-10;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAQFSKTAAKGEAAERPGAAVA 24  
|||||  
Db 1 GAQFSKTAAKGEAAERPGAAVA 24  
|||||

RESULT 2  
US-10-802-644-1  
; Sequence 1, Application US/10802644

```
; Publication No. US20040180836A1
; GENERAL INFORMATION:
; APPLICANT: Martin, Linda
; APPLICANT: Adler, Kenneth
; APPLICANT: Li, Yuehua
; TITLE OF INVENTION: BLOCKING PEPTIDE FOR INFLAMMATORY CELL SECRETION
; FILE REFERENCE: 5051.574
; CURRENT APPLICATION NUMBER: US/10/802,644
; CURRENT FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: US/10/180,753
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
; NAME/KEY: MOD RES
; LOCATION: (1)-(1)
; OTHER INFORMATION: N-terminal myristoylation
;
US-10-802-644-1

Query Match 100.0%; Score 115; DB 16; Length 24;
Best Local Similarity 100.0%; Pred. No. 8.6e-10; Indels 0; Gaps 0;
Matches 24; Conservative 0; Mismatches 0;

QY 1 GAQFSKTAAKGEAAARPGEAAVA 24
||| ||||| ||||| ||||| |||||
Db 1 GAQFSKTAAKGEAAARPGEAAVA 24

RESULT 3
US-09-864-761-44320
; Sequence 44320, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44320
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004845.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.85
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
;
US-09-864-761-44320

Query Match 33.9%; Score 39; DB 9; Length 24;
Best Local Similarity 50.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 7;

QY 4 FSKTAAKGEAAARPG 19
||| ||| ||| |||
Db 4 FSHVPANSTAAGEKPG 19

RESULT 4
US-10-390-472-35
; Sequence 35, Application US/10390472
; Publication No. US20040087013A1
; GENERAL INFORMATION:
; APPLICANT: Holmes, Todd
; APPLICANT: Zhang, Shuguang
; APPLICANT: Rich, Alexander
; APPLICANT: DiPersio, C. Michael
; APPLICANT: Lockshin, Curtis
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; THEREFOR
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/390,472
; FILING DATE: 17-Mar-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/293,284
; FILING DATE: 22-AUG-1994
; APPLICATION NUMBER: 07/973,326
; FILING DATE: 28-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
```

```
;
;
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: MIT-6008A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 35:
US-10-390-472-35

Query Match          32.2%; Score 37; DB 16; Length 16;
Best Local Similarity 56.2%; Pred. No. 1.3e+02;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      6 KTAAKGEAAARPEGA 21
Db      1 KAEAKAEAKAEAKAE 16

RESULT 5
US-09-939-537-52
; Sequence 52, Application US/09939537
; Publication No. US20030138410A1
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; Banapour, Babak
; Romeo, Charles
; Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,537
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
```

```
;
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-09-939-537-52

Query Match          32.2%; Score 37; DB 10; Length 20;
Best Local Similarity 56.2%; Pred. No. 1.6e+02;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      6 KTAAKGEAAARPEGA 21
Db      3 KAEAKAEAKAEAKAE 18

RESULT 6
US-10-372-876-507
; Sequence 507, Application US/10372876
; Publication No. US20030204071A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A. et al.
; TITLE OF INVENTION: 110 Human Secreted Proteins
; FILE REFERENCE: P2021P1
; CURRENT APPLICATION NUMBER: US/10/372,876
; CURRENT FILING DATE: 2003-02-26
; PRIOR APPLICATION NUMBER: 09/334,595
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: PCT/US98/27059
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,007
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,057
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,006
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,369
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,367
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,169
; PRIOR FILING DATE: 1997-12-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 507
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-372-876-507

Query Match          31.3%; Score 36; DB 12; Length 20;
Best Local Similarity 70.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      13 AAARPEGEAA 22
Db      8 AQAEKPEGEA 17

RESULT 7
US-10-097-065-507
; Sequence 507, Application US/10097065
; Publication No. US20030055236A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A. et al.
; TITLE OF INVENTION: 110 Human Secreted Proteins
; FILE REFERENCE: P2021P1
; CURRENT APPLICATION NUMBER: US/10/097,065
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: PCT/US98/27059
```

; PRIOR FILING DATE: 1998-12-17  
; PRIOR APPLICATION NUMBER: 60/070,923  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/068,007  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/068,057  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/068,006  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/068,369  
; PRIOR FILING DATE: 1997-12-19  
; PRIOR APPLICATION NUMBER: 60/068,367  
; PRIOR FILING DATE: 1997-12-19  
; PRIOR APPLICATION NUMBER: 60/068,368  
; PRIOR FILING DATE: 1997-12-19  
; PRIOR APPLICATION NUMBER: 60/068,169  
; PRIOR FILING DATE: 1997-12-19  
; PRIOR APPLICATION NUMBER: 60/068,053  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/068,064  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/068,054  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/068,008  
; PRIOR FILING DATE: 1997-12-18  
; PRIOR APPLICATION NUMBER: 60/068,365  
; PRIOR FILING DATE: 1997-12-19  
; NUMBER OF SEQ ID NOS: 672  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 507  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-097-065-507

Query Match 31.3%; Score 36; DB 14; Length 20;  
Best Local Similarity 70.0%; Pred. No. 2.3e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 13 AAARPGEEA 22  
|:|:|:|:|  
DB 8 AQAEPGEA 17

RESULT 8  
US-10-109-048-902  
; Sequence 902, Application US/10109048  
; Publication No. US20040107461A1  
; GENERAL INFORMATION:  
; APPLICANT: COMMURI, PADMA  
; APPLICANT: KEELING, PETER L.  
; APPLICANT: RAMIREZ, NONA  
; APPLICANT: MCKEAN, ANGELA  
; APPLICANT: GAO, ZHONG  
; APPLICANT: GUAN, HANPING  
; TITLE OF INVENTION: GLUCAN CHAIN LENGTH DOMAINS  
; FILE REFERENCE: 2461-76  
; CURRENT APPLICATION NUMBER: US/10/109,048  
; CURRENT FILING DATE: 2003-03-04  
; PRIOR APPLICATION NUMBER: 60/279,720  
; PRIOR FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 1154  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 902  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Accession No. 15903076  
US-10-109-048-902

Query Match 30.4%; Score 35; DB 16; Length 21;  
Best Local Similarity 58.3%; Pred. No. 3.3e+02;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 GAQFSKTAAKGE 12  
|:|:|:|:|  
DB 10 GAFFSKTGGLGD 21  
RESULT 9  
US-10-109-048-904  
; Sequence 904, Application US/10109048  
; Publication No. US20040107461A1  
; GENERAL INFORMATION:  
; APPLICANT: COMMURI, PADMA  
; APPLICANT: KEELING, PETER L.  
; APPLICANT: RAMIREZ, NONA  
; APPLICANT: MCKEAN, ANGELA  
; APPLICANT: GAO, ZHONG  
; APPLICANT: GUAN, HANPING  
; TITLE OF INVENTION: GLUCAN CHAIN LENGTH DOMAINS  
; FILE REFERENCE: 2461-76  
; CURRENT APPLICATION NUMBER: US/10/109,048  
; CURRENT FILING DATE: 2003-03-04  
; PRIOR APPLICATION NUMBER: 60/279,720  
; PRIOR FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 1154  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 904  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown Organism: Accession No. 15900991  
US-10-109-048-904

Query Match 30.4%; Score 35; DB 16; Length 21;  
Best Local Similarity 58.3%; Pred. No. 3.3e+02;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 GAQFSKTAAKGE 12  
|:|:|:|:|  
DB 10 GAFFSKTGGLGD 21

RESULT 10  
US-10-390-472-33  
; Sequence 33, Application US/10390472  
; Publication No. US20040087013A1  
; GENERAL INFORMATION:  
; APPLICANT: Holmes, Todd  
; APPLICANT: Zhang, Shuguang  
; APPLICANT: Rich, Alexander  
; APPLICANT: DiPersio, C. Michael  
; APPLICANT: Lockshin, Curtis  
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY  
; SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES  
; THEREFOR  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02173-4799  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/390,472  
; FILING DATE: 17-Mar-2003  
; CLASSIFICATION: 435



;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/293,284  
;; FILING DATE: 22-AUG-1994  
;; APPLICATION NUMBER: 07/973,326  
;; FILING DATE: 28-DEC-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Brook, David E.  
;; REGISTRATION NUMBER: 22,592  
;; REFERENCE/DOCKET NUMBER: MIT-6008A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 861-6240  
;; TELEFAX: (617) 861-9540  
;; INFORMATION FOR SEQ ID NO: 33:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 16 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 33:  
US-10-390-472-33

Query Match 29.6%; Score 34; DB 16; Length 16;  
Best Local Similarity 50.0%; Pred. No. 3.5e+02;  
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 SKTAAGKEAAERPG 20  
Db 1 AKAAKAAKAAKAAE 16  
|:::| | | | |

RESULT 11  
US-09-765-301-28  
; Sequence 28, Application US/09765301  
; Patent No. US20020037848A1  
; GENERAL INFORMATION:  
; APPLICANT: EISENBACH-SCHWARTZ, Michal  
; APPLICANT: YOLLES, Etti  
; TITLE OF INVENTION: THE USE OF COPOLYMER 1 AND RELATED PEPTIDES AND POLYPEPTIDES AND  
; FILE REFERENCE: EIS-SCHWARTZ18  
; CURRENT APPLICATION NUMBER: US/09/765,301  
; PRIOR FILING DATE: 2001-01-22  
; PRIOR APPLICATION NUMBER: 06/209,799  
; PRIOR FILING DATE: 2000-06-07  
; PRIOR APPLICATION NUMBER: 09/620,216  
; PRIOR FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 28  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
US-09-765-301-28

Query Match 28.7%; Score 33; DB 9; Length 15;  
Best Local Similarity 50.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 AQFSKTAAGKEAAA 15  
Db 1 AEYAKAAAAAAA 14  
|:::| | | | |

RESULT 12  
US-09-765-644A-28  
; Sequence 28, Application US/09765644A  
; Publication No. US20030004099A1  
; GENERAL INFORMATION:  
; APPLICANT: EISENBACH-SCHWARTZ, Michal  
; APPLICANT: COHEN, Irun R.

;; APPLICANT: SELA, Michael  
;; APPLICANT: YOLLES, Etti  
;; APPLICANT: KIPNIS, Jonathan  
;; TITLE OF INVENTION: THE USE OF COPOLYMER 1 AND RELATED PEPTIDES AND POLYPEPTIDES AND  
;; FILE REFERENCE: EIS-SCHWARTZ13B  
;; CURRENT APPLICATION NUMBER: US/09/765,644A  
;; CURRENT FILING DATE: 2001-01-22  
;; PRIOR APPLICATION NUMBER: 09/487,793  
;; PRIOR FILING DATE: 2000-01-20  
;; PRIOR APPLICATION NUMBER: 06/209,799  
;; PRIOR FILING DATE: 2000-06-07  
;; PRIOR APPLICATION NUMBER: 09/620,216  
;; PRIOR FILING DATE: 2000-07-20  
;; NUMBER OF SEQ ID NOS: 33  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 28  
;; LENGTH: 15  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Synthetic construct  
US-09-765-644A-28

Query Match 28.7%; Score 33; DB 10; Length 15;  
Best Local Similarity 50.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 AQFSKTAAGKEAAA 15  
Db 1 AEYAKAAAAAAA 14  
|:::| | | | |

RESULT 13  
US-10-438-538-32  
; Sequence 32, Application US/10438538  
; Publication No. US20040006022A1  
; GENERAL INFORMATION:  
; APPLICANT: Strominger, Jack L.  
; APPLICANT: Fridkis-Hareli, Masha  
; TITLE OF INVENTION: Synthetic Peptides and Methods of use for Autoimmune  
; FILE REFERENCE: 24655-013DIV2  
; CURRENT APPLICATION NUMBER: US/10/438,538  
; CURRENT FILING DATE: 2003-05-15  
; PRIOR APPLICATION NUMBER: 09/359,099  
; PRIOR FILING DATE: 1999-07-22  
; PRIOR APPLICATION NUMBER: 60/093,859  
; PRIOR FILING DATE: 1998-07-23  
; PRIOR APPLICATION NUMBER: 60/123,675  
; PRIOR FILING DATE: 1999-03-09  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 32  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: peptide of predetermined sequence for testing of  
; OTHER INFORMATION: activity in MHC Class II assays.  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1)..(15)  
US-10-438-538-32

Query Match 28.7%; Score 33; DB 15; Length 15;  
Best Local Similarity 50.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 AQFSKTAAGKEAAA 15  
Db 1 AEYAKAAAAAAA 14  
|:::| | | | |

RESULT 14  
US-09-778-200-6  
; Sequence 6, Application US/09778200  
; Patent No. US20020160471A1  
; GENERAL INFORMATION:  
; APPLICANT: Kisdai, John  
; APPLICANT: Grodzinsky, Alan  
; APPLICANT: Zhang, Shuguang  
; TITLE OF INVENTION: Peptide Scaffold Encapsulation of Tissue  
; TITLE OF INVENTION: Cells and Uses Thereof  
; FILE REFERENCE: 01997/537001  
; CURRENT APPLICATION NUMBER: US/09/778,200  
; CURRENT FILING DATE: 2000-02-06  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Designed Peptide  
US-09-778-200-6

Query Match 28.7%; Score 33; DB 9; Length 16;  
Best Local Similarity 61.5%; Pred. No. 4.8e+02;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 9 AKGEAAERPGEA 21  
Db 3 AKAEAKAEAEAE 15

RESULT 15  
US-09-778-200-21  
; Sequence 21, Application US/09778200  
; Patent No. US20020160471A1  
; GENERAL INFORMATION:  
; APPLICANT: Kisdai, John  
; APPLICANT: Grodzinsky, Alan  
; APPLICANT: Zhang, Shuguang  
; TITLE OF INVENTION: Peptide Scaffold Encapsulation of Tissue  
; TITLE OF INVENTION: Cells and Uses Thereof  
; FILE REFERENCE: 01997/537001  
; CURRENT APPLICATION NUMBER: US/09/778,200  
; CURRENT FILING DATE: 2000-02-06  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 21  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Designed Peptide  
US-09-778-200-21

Query Match 28.7%; Score 33; DB 9; Length 16;  
Best Local Similarity 50.0%; Pred. No. 4.8e+02;  
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
QY 6 KTAKEAAERPGEA 21  
Db 1 KAKAKAEAEAEAE 16

Search completed: October 7, 2004, 18:10:46  
Job time : 85.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:50:43 ; Search time 22 Seconds  
(without alignments)  
56.319 Million cell updates/sec

Title: US-10-802-644-1

Perfect score: 115

Sequence: 1 GAQFSKTAAGBAARPGBAVA 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 189146

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	32.2	16	1	US-08-346-849-35
2	37	32.2	16	2	US-08-293-284A-35
3	37	32.2	16	4	US-08-898-300-35
4	37	32.2	19	2	US-08-726-306A-72
5	37	32.2	20	2	US-08-284-391B-52
6	37	32.2	20	3	US-09-218-950-52
7	34	29.6	16	1	US-08-346-849-33
8	34	29.6	16	2	US-08-293-284A-33
9	34	29.6	16	4	US-08-898-300-33
10	34	29.6	21	2	US-08-557-309B-58
11	34	29.6	21	3	US-08-929-414-11
12	34	29.6	21	3	US-08-834-306-56
13	34	29.6	21	3	US-08-834-306-58
14	34	29.6	21	3	US-08-993-674A-56
15	34	29.6	21	3	US-08-993-674A-58
16	34	29.6	21	4	US-09-256-976-56
17	34	29.6	21	4	US-09-256-976-58
18	33	28.7	16	1	US-08-346-849-34
19	33	28.7	16	1	US-08-346-849-36
20	33	28.7	16	2	US-08-293-284A-34
21	33	28.7	16	2	US-08-293-284A-36
22	33	28.7	16	4	US-08-898-300-34
23	33	28.7	16	4	US-08-898-300-36
24	32	27.8	21	1	US-08-127-499A-33
25	32	27.8	21	1	US-08-482-847-33
26	32	27.8	22	3	US-08-990-015-5
27	32	27.8	22	4	US-09-589-978C-7

28	31	27.0	17	4	US-09-140-749-14	Sequence 14, Appl
29	31	27.0	18	4	US-09-005-167A-99	Sequence 99, Appl
30	31	27.0	21	2	US-08-997-080-6	Sequence 6, Appl
31	31	27.0	21	2	US-08-997-362-6	Sequence 6, Appl
32	31	27.0	21	3	US-08-873-970-6	Sequence 6, Appl
33	31	27.0	21	3	US-09-095-855-6	Sequence 6, Appl
34	31	27.0	21	3	US-08-705-347A-6	Sequence 6, Appl
35	31	27.0	21	3	US-09-143-124-22	Sequence 22, Appl
36	31	27.0	21	4	US-09-324-542-6	Sequence 6, Appl
37	31	27.0	21	4	US-09-205-426-6	Sequence 6, Appl
38	31	27.0	21	4	US-09-200-643-6	Sequence 6, Appl
39	31	27.0	22	4	US-07-946-180B-16	Sequence 16, Appl
40	30	26.1	16	1	US-08-346-849-43	Sequence 43, Appl
41	30	26.1	16	2	US-08-293-284A-43	Sequence 43, Appl
42	30	26.1	16	4	US-09-230-548-17	Sequence 17, Appl
43	30	26.1	16	4	US-08-898-300-43	Sequence 43, Appl
44	30	26.1	20	2	US-08-747-137-61	Sequence 61, Appl
45	30	26.1	21	2	US-08-557-309B-60	Sequence 60, Appl

## ALIGNMENTS

RESULT 1  
US-08-346-849-35  
; Sequence 35, Application US/08346849  
; Patent No. 5670483  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Shuguang  
; APPLICANT: Lockshin, Curtis  
; APPLICANT: Rich, Alexander  
; APPLICANT: Holmes, Todd  
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY  
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES  
; TITLE OF INVENTION: THEREFOR  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02173-4799  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/346,849  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,326  
; FILING DATE: 28 DECEMBER 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brook, David E.  
; REGISTRATION NUMBER: 22,592  
; REFERENCE/DOCKET NUMBER: MIT-6008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 861-6240  
; TELEFAX: (617) 861-9540  
; INFORMATION FOR SEQ ID NO: 35:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-346-849-35

Query Match 32.2%; Score 37; DB 1; Length 16;  
Best Local Similarity 56.2%; Pred. No. 17;  
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
QY      6 KTAAKGEAAARPGEA 21
      |||||
Db      1 KAEAKAEAKAEAKAEA 16

RESULT 2
US-08-293-284A-35
; Sequence 35, Application US/08293284A
; Patent No. 5955343
; GENERAL INFORMATION:
; APPLICANT: Holmes, Todd
; APPLICANT: Zhang, Shuguang
; APPLICANT: Rich, Alexander
; APPLICANT: DiPersio, C. Michael
; APPLICANT: Lockshin, Curtis
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/293,284A
; FILING DATE: 22-AUG-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: 07/973,326
; FILING DATE: 28-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: MIT-6008A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-293-284A-35

Query Match      32.2%; Score 37; DB 2; Length 16;
Best Local Similarity 56.2%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      6 KTAAKGEAAARPGEA 21
      |||||
Db      1 KAEAKAEAKAEAKAEA 16

RESULT 3
US-08-898-300-35
; Sequence 35, Application US/08898300
; Patent No. 6548630
; GENERAL INFORMATION:
; APPLICANT: Zhang, Shuguang
; APPLICANT: Lockshin, Curtis
; APPLICANT: Rich, Alexander
; APPLICANT: Holmes, Todd
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY

; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173-4799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 08/346,849
; FILING DATE: 30 NOVEMBER 1994
; PRIOR APPLICATION NUMBER: 07/973,326
; FILING DATE: 28 DECEMBER 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: MIT-6008PB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 861-6240
; TELEFAX: (781) 861-9540
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-898-300-35

Query Match      32.2%; Score 37; DB 4; Length 16;
Best Local Similarity 56.2%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY      6 KTAAKGEAAARPGEA 21
      |||||
Db      1 KAEAKAEAKAEAKAEA 16

RESULT 4
US-08-726-306A-72
; Sequence 72, Application US/08726306A
; Patent No. 5958684
; GENERAL INFORMATION:
; APPLICANT: van Leeuwen, Frederik Willem
; APPLICANT: Burbach, Johannes Peter Henri
; APPLICANT: Grosveld, Franklin G.
; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
; NUMBER OF SEQUENCES: 189
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1 Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,306A
```

```
; FILING DATE: 02-Oct-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 95/20080.4
; FILING DATE: 02-Oct-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,832
; FILING DATE: 01-Jan-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-726-306A-72

Query Match 32.2%; Score 37; DB 2; Length 19;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 SKTAAKGEEAAERP 18
|:|:|:|:|:|
Db 2 SRTGASGAASSARP 15

RESULT 5
US-08-284-391B-52
; Sequence 52, Application US/08284391B
; Patent No. 5851828
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391B
; FILING DATE: 02-AUG-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; FILING DATE: 07-MAR-1991
; NAME: Elbing, Karen L
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
```

```
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-284-391B-52

Query Match 32.2%; Score 37; DB 2; Length 20;
Best Local Similarity 56.2%; Pred. No. 22;
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 KTAAKGEAAERPGEA 21
|:|:|:|:|:|
Db 3 KAEAKAEAKAEAKAE 18

RESULT 6
US-09-218-950-52
; Sequence 52, Application US/09218950
; Patent No. 6284240
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/218,950
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-218-950-52
```

Query Match 32.2%; Score 37; DB 3; Length 20;  
Best Local Similarity 56.2%; Pred. No. 22;  
Matches 9; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 KTAAKGEAAARPEA 21  
| | | | | | | | | |  
DB 3 KAAKAAKAAKAAKAE 18

RESULT 7  
US-08-346-849-33  
; Sequence 33, Application US/08346849  
; Patent No. 5670483  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Shuguang  
; APPLICANT: Lockshin, Curtis  
; APPLICANT: Rich, Alexander  
; APPLICANT: Holmes, Todd  
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY  
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02173-4799  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/346,849  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,326  
; FILING DATE: 28 DECEMBER 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brook, David E.  
; REGISTRATION NUMBER: 22,592  
; REFERENCE/DOCKET NUMBER: MIT-6008  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 861-6240  
; TELEFAX: (617) 861-9540  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-346-849-33

Query Match 29.6%; Score 34; DB 1; Length 16;  
Best Local Similarity 50.0%; Pred. No. 50;  
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 SKTAAKGEAAARPE 20  
: | | | | | | | | | |  
DB 1 AKAAKAAKAAKAAKAE 16

RESULT 8  
US-08-293-284A-33  
; Sequence 33, Application US/08293284A  
; Patent No. 5955343  
; GENERAL INFORMATION:  
; APPLICANT: Holmes, Todd  
; APPLICANT: Zhang, Shuguang

; APPLICANT: Rich, Alexander  
; APPLICANT: DiPersio, C. Michael  
; APPLICANT: Lockshin, Curtis  
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY  
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02173-4799  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/293,284A  
; FILING DATE: 22-AUG-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,326  
; FILING DATE: 28-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brook, David E.  
; REGISTRATION NUMBER: 22,592  
; REFERENCE/DOCKET NUMBER: MIT-6008A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 861-6240  
; TELEFAX: (617) 861-9540  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 16 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-293-284A-33

Query Match 29.6%; Score 34; DB 2; Length 16;  
Best Local Similarity 50.0%; Pred. No. 50;  
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 SKTAAKGEAAARPE 20  
: | | | | | | | | | |  
DB 1 AKAAKAAKAAKAAKAE 16

RESULT 9  
US-08-898-300-33  
; Sequence 33, Application US/08898300  
; Patent No. 6548630  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Shuguang  
; APPLICANT: Lockshin, Curtis  
; APPLICANT: Rich, Alexander  
; APPLICANT: Holmes, Todd  
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY  
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02173-4799  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/898,300  
FILING DATE: 22 JULY 1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/346,849  
FILING DATE: 30 NOVEMBER 1994  
PRIOR APPLICATION DATA: 07/973,326  
APPLICATION NUMBER: 07/973,326  
FILING DATE: 28 DECEMBER 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: MIT-6008FB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 861-6240  
TELEFAX: (781) 861-9540  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-898-300-33

Query Match 29.6%; Score 34; DB 4; Length 16;  
Best Local Similarity 50.0%; Pred. No. 50;  
Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 SKTAAGGAAAEPRGE 20  
DB 1 AXAEAKAEAKAEAKAE 16

## RESULT 10

US-08-557-309B-58  
Sequence 58, Application US/08557309B  
Patent No. 5916572

GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Houghton, Raymond L.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T  
NUMBER OF SEQUENCES: 69  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/557,309B  
FILING DATE: 14-NOV-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.422  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids

TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-557-309B-58

Query Match 29.6%; Score 34; DB 2; Length 21;  
Best Local Similarity 47.4%; Pred. No. 67;  
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTAAGGAAAEPRGEAAVA 24  
DB 1 KAAAPAKAAAPAKAAAA 19

## RESULT 11

US-08-929-414-11  
Sequence 11, Application US/08929414  
Patent No. 5942403

GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Haughton, Raymond  
APPLICANT: Skeiky, Yasir A.W.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION  
OF T. CRUZI INFECTION  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/929,414  
FILING DATE: 15-SEP-1997  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.406C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-929-414-11

Query Match 29.6%; Score 34; DB 2; Length 21;  
Best Local Similarity 47.4%; Pred. No. 67;  
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTAAGGAAAEPRGEAAVA 24  
DB 1 KAAAPAKAAAPAKAAAA 19

## RESULT 12

US-08-834-306-56  
Sequence 56, Application US/08834306  
Patent No. 6054135

GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Lodes, Michael J.

```

; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,306
; FILING DATE: 15-APR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
; US-08-834-306-56

Query Match 29.6%; Score 34; DB 3; Length 21;
Best Local Similarity 47.4%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTAAKGEAAARPGEAAVA 24
DB 1 KAAAPAKAAAPAKAAAA 19

RESULT 13
US-08-834-306-58
; Sequence 58, Application US/08834306
; Patent No. 6054135
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,306
; FILING DATE: 15-APR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
; US-08-834-306-56

Query Match 29.6%; Score 34; DB 3; Length 21;
Best Local Similarity 47.4%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTAAKGEAAARPGEAAVA 24
DB 1 KAAAPAKAAAPAKAAAA 19

RESULT 13
US-08-834-306-58
; Sequence 58, Application US/08834306
; Patent No. 6054135
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,306
; FILING DATE: 15-APR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
; US-08-834-306-58

Query Match 29.6%; Score 34; DB 3; Length 21;
Best Local Similarity 47.4%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTAAKGEAAARPGEAAVA 24
DB 1 KAAAPAKAAAPAKAAAA 19

RESULT 14
US-08-993-674A-56
; Sequence 56, Application US/08993674A
; Patent No. 6228372
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Smith, John M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,674A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
; US-08-993-674A-56

Query Match 29.6%; Score 34; DB 3; Length 21;
Best Local Similarity 47.4%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTAAKGEAAARPGEAAVA 24
DB 1 KAAAPAKAAAPAKAAAA 19
```

```

; REFERENCE/DOCKET NUMBER: 210121.422C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
; US-08-834-306-58

Query Match 29.6%; Score 34; DB 3; Length 21;
Best Local Similarity 47.4%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTAAKGEAAARPGEAAVA 24
DB 1 KAAAPAKAAAPAKAAAA 19

RESULT 14
US-08-993-674A-56
; Sequence 56, Application US/08993674A
; Patent No. 6228372
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Smith, John M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,674A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
; US-08-993-674A-56

Query Match 29.6%; Score 34; DB 3; Length 21;
Best Local Similarity 47.4%; Pred. No. 67;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 6 KTAAKGEAAARPGEAAVA 24
DB 1 KAAAPAKAAAPAKAAAA 19
```



```

RESULT 15
US-08-993-674A-58
; Sequence 58, Application US/08993674A
; Patent No. 6228372
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Smith, John M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/993,674A
; APPLICATION NUMBER: US/08/993,674A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-993-674A-58

```

Query Match 29.6%; Score 34; DB 3; Length 21;  
Best Local Similarity 47.4%; Pred. NO. 67;  
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 6 KTAAKGEAAAEERPGEAAVA 24  
||| ||| :|||  
Db 1 KAAAPAKAAAAAPAKAAAA 19

Search completed: October 7, 2004, 17:58:13  
Job time : 23 secs

**This Page Blank (uspto)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:47:47 ; Search time 23.5 seconds  
(without alignments)  
98.238 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAAEGAGAEVKRASAEAKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 4740

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Match	Length	ID	Description
1	37	32.2	24	2	S33302	hydrogenase (EC 1.
2	36	31.3	20	2	S38291	30K allergen - vel
3	36	31.3	22	2	D47256	kinetoplast DNA-as
4	32	27.8	23	2	I64839	SP-A1 (alpha, beta
5	30.5	26.5	20	2	S38294	32K allergen - com
6	30	26.1	24	2	A61505	pollen allergen Ph
7	30	26.1	24	2	F45357	Kex2/subtilisin-li
8	30	26.1	24	2	C45357	Kex2/subtilisin-li
9	29	25.2	20	2	S19618	globin - polychaet
10	28	24.3	15	2	PQ0692	photosystem I 18.5
11	28	24.3	16	2	S38292	30K allergen - rye
12	28	24.3	21	2	S32883	DNA gyrase chain A
13	26.5	23.0	22	2	A26376	peptidyl-dipeptida
14	26	22.6	20	2	PH0918	T-cell receptor be
15	26	22.6	15	2	PS0260	56K protein 2402 -
16	26	22.6	17	2	C30221	histone H2A.8 - ch
17	26	22.6	18	2	S45373	translation elonga
18	26	22.6	20	2	PQ0687	photosystem I 14.1
19	25	21.7	16	2	A11488	taurocyamine kinas
20	25	21.7	20	2	S32502	calpain (EC 3.4.22
21	24.5	21.3	19	2	B29501	fibrinopeptide A -
22	24	20.9	13	2	S23638	Ig kappa chain J s
23	24	20.9	20	2	S06149	photosystem I chai
24	24	20.9	20	2	S38288	50K allergen - per
25	24	20.9	20	2	PC1150	equinotoxin 1B - s
26	23	20.0	10	2	H60588	sperm-activating p
27	23	20.0	12	1	A43975	locustamyotropin -
28	23	20.0	12	2	PH1481	T-cell receptor be
29	23	20.0	14	2	S23639	Ig kappa chain J s

30	23	20.0	14	2	E81280	probable proteolys
31	23	20.0	15	2	PA0087	cytochrome c2 - fu
32	23	20.0	15	2	A49177	22K protein p1, mi
33	23	20.0	16	2	PA0103	L-lactate dehydrog
34	23	20.0	19	2	H27480	hydrogenase (EC 1.
35	23	20.0	20	2	S33787	pancreatic elastas
36	23	20.0	20	2	B44920	2-halobenzoate 1,2
37	23	20.0	20	2	A58903	metalloprotease
38	23	20.0	22	2	S68616	histone H1 - sea u
39	23	20.0	22	2	PC2134	maltose transport
40	23	20.0	23	2	S38991	glycine reductase
41	23	20.0	24	2	S70351	peptidyl-dipeptida
42	22	19.1	12	2	A20907	Ig kappa chain J1
43	22	19.1	13	2	B58533	CD61 homolog - cha
44	22	19.1	13	2	B20907	Ig kappa-1 chain J
45	22	19.1	13	2	B25448	Ig kappa-1 chain,

#### ALIGNMENTS

##### RESULT 1

S33302

hydrogenase (EC 1.18.99.1) - Chlamydomonas reinhardtii (fragment)

C:Species: Chlamydomonas reinhardtii

C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 30-Sep-1993

C:Accession: S33302

R:Happe, T.; Naber, J.D.

Eur. J. Biochem. 214, 475-481, 1993

A:Title: Isolation, characterization and N-terminal amino acid sequence of hydrogenase fi

A:Reference number: S33302; MUID:93292509; PMID:8513797

A:Accession: S33302

A:Molecule type: protein

A:Residues: 1-24 <HAP>

C:Keywords: chloroplast; oxidoreductase

Query Match 32.2%; Score 37; DB 2; Length 24;

Best Local Similarity 52.9%; Pred. No. 1.1e+02;

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVKRASAE 19

DB 2 APAAEAPLSHVQQAALAE 18

##### RESULT 2

S38291

30K allergen - velvet grass (fragment)

C:Species: Holcus lanatus (velvet grass)

C:Date: 19-Mar-1997 #sequence\_revision 06-Jun-1997 #text\_change 19-May-2000

C:Accession: S38291

R:Petersen, A.; Schramm, G.; Becker, W.M.; Schlaak, M.

Biol. Chem. Hoppe-Seyler 374, 855-861, 1993

A:Title: Comparison of four grass pollen species concerning their allergens of grass gro

A:Reference number: S38288; MUID:94092339; PMID:7505588

A:Accession: S38291

A:Molecule type: protein

A:Residues: 1-20 <PET>

Query Match 31.3%; Score 36; DB 2; Length 20;

Best Local Similarity 70.0%; Pred. No. 1.3e+02;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 APAAEGAGAE 12

DB 9 APAATGAGGD 18

##### RESULT 3

D47256

kinetoplast DNA-associated protein p15 - Crithidia fasciculata (fragment)

C:Species: Crithidia fasciculata

C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Nov-1994

C:Accession: D47256  
R:Yu, C.; Ray, D.S.

Proc. Natl. Acad. Sci. U.S.A. 90, 1786-1789, 1993

A:Title: Isolation of proteins associated with kinetoplast DNA networks in vivo.

A:Reference number: A47256; MUID:93189582; PMID:8446592

C:Accession: D47256

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-22 <XU1>

A:Note: sequence extracted from NCBI backbone (NCBIP:126909)

Query Match 31.3%; Score 36; DB 2; Length 22;

Best Local Similarity 42.1%; Pred. No. 1.4e+02;

Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 3 APAERGAGAEVKRAAEAK 21

|||||:|:|:|:|

Db 4 APAKKAAPKAAKASTPAK 22

RESULT 4

I64839

SP-A1 (alpha, beta, epsilon) and SP-A2 (alpha, beta) - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 29-May-1998 #sequence revision 29-May-1998 #text\_change 20-Aug-1999

C:Accession: I64839; I64842; I64845; I64838

R:McCormick, S.M.; Boggaram, V.; Mendelson, C.R.

Am. J. Physiol. 266, L354-L366, 1994

A:Title: Characterization of mRNA transcripts and organization of human SP-A1 and SP-A2

A:Reference number: I51910; MUID:94234365; PMID:8179012

C:Accession: I64839

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-23 <RES>

A:Cross-references: GB:S69681; NID:g546673; PIDN:AAB30731.1; PID:g546674

A:Accession: I64842

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-23 <RE2>

A:Cross-references: GB:S69686; NID:g546681; PIDN:AAB30735.1; PID:g546682

A:Accession: I64841

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Cross-references: GB:S69685; NID:g546679; PIDN:AAB30734.1; PID:g546680

A:Accession: I64845

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-23 <RE4>

A:Cross-references: GB:S69690; NID:g546687; PIDN:AAB30738.1; PID:g546688

A:Accession: I64838

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-23 <RE5>

A:Cross-references: GB:S69680; NID:g546671; PIDN:AAB30730.1; PID:g546672

C:Genetics:

A:Gene: SP-A2; SP-A1

C:Superfamily: mannose-binding lectin; C-type lectin homology

Query Match 27.8%; Score 32; DB 2; Length 23;

Best Local Similarity 70.0%; Pred. No. 5.3e+02;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 AAEAGAEVK 14

|||||

Db 14 AASGAACEVK 23

RESULT 5

S38294

32k allergen - common timothy (fragment)

C:Species: Phleum pratense (common timothy)

C:Date: 19-May-1994 #sequence\_revision 27-Feb-1997 #text\_change 07-May-1999

C:Accession: S38294

R:Petersen, A.; Schramm, G.; Becker, W.M.; Schlaak, M.

Biol. Chem. Hoppe-Seyler 374, 855-861, 1993

A:Title: Comparison of four grass pollen species concerning their allergens of grass

A:Reference number: S38288; MUID:94092339; PMID:7505588

C:Accession: S38294

A:Molecule type: protein

A:Residues: 1-20 <PET>

Query Match 26.5%; Score 30.5; DB 2; Length 20;

Best Local Similarity 52.9%; Pred. No. 7.5e+02;

Matches 9; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 1 GTAPAA-EGAGAEVKRA 16

|||||:|:|:|:|

Db 4 GYAPATPAAGAEAGKA 20

RESULT 6

A61505

pollen allergen Phl p V - common timothy (fragment)

C:Species: Phleum pratense (common timothy)

C:Date: 15-Oct-1994 #sequence\_revision 15-Oct-1994 #text\_change 07-May-1999

C:Accession: A61505

R:Matthiesen, F.; Lowenstein, H.

Clin. Exp. Allergy 21, 297-307, 1991

A:Title: Group V allergens in grass pollens. I. Purification and characterization of the

A:Reference number: A61505; MUID:91322736; PMID:1863892

C:Accession: A61505

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-24 <MAT>

A:Note: 2-Ala, 5-Gly, and 15-Glu were also found

C:Superfamily: grass pollen allergen IX

C:Keywords: hydroxyproline; pollen

F:6,9,12,18,21,24/Modified site: hydroxyproline (Pro) #status experimental

Query Match 26.1%; Score 30; DB 2; Length 24;

Best Local Similarity 85.7%; Pred. No. 1e+03;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAEEGAG 10

|||||

Db 9 PAAPGAG 15

RESULT 7

F45357

Kex2/subtilisin-like proprotein convertase PC4-C - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Nov-1994

C:Accession: F45357

R:Seidah, N.G.; Day, R.; Hamelin, J.; Gaspar, A.; Collard, M.W.; Chretien, M.

Mol. Endocrinol. 6, 1559-1570, 1992

A:Title: Testicular expression of PC4 in the rat: molecular diversity of a novel germ cell

A:Reference number: A45357; MUID:93078790; PMID:1448111

C:Accession: F45357

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-24 <SEI>

A:Note: sequence extracted from NCBI backbone (NCBIP:118886)

Query Match 26.1%; Score 30; DB 2; Length 24;

Best Local Similarity 31.2%; Pred. No. 1e+03;

Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 TAPAEAGAEVKRAS 17

::::|:|:|:|

Db 8 SSPASSGGGSTATSS 23

RESULT 8

C45357

```

Kex2/subtilisin-like proprotein convertase PC4-C - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C>Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C:Accession: C45357
R:Seidah, N.G.; Day, R.; Hamelin, J.; Gaspar, A.; Collard, M.W.; Chretien, M.
Mol. Endocrinol. 6, 1559-1570, 1992
A:Title: Testicular expression of PC4 in the rat: molecular diversity of a novel germ cell
A:Reference number: A45357; MUID:93078790; PMID:1448111
A:Accession: C45357
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-24 <SEI>
A>Note: sequence extracted from NCBI backbone (NCBIP:118879)

Query Match      26.1%; Score 30; DB 2; Length 24;
Best Local Similarity 31.2%; Pred. No. 1e+03;
Matches 5; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY  2  TAPAAEGAGAEVKRAS 17
    :|:|:|:|:|:|:|:|:|:
Db   8  SSPASSGGGGSTATHSS 23

RESULT 9
S19618
globin - polychaete (Eudistylia vancouveri) (fragment)
N:Alternate names: chlorocruorin
C:Species: Eudistylia vancouveri
C>Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 05-Dec-1998
C:Accession: S19618
R:Qabar, A.N.; Stern, M.S.; Walz, D.A.; Chiu, J.T.; Timkovich, R.; Wall, J.S.; Kapp, O.H.
J. Mol. Biol. 222, 1109-1129, 1991
A:Title: Hierarchy of globin complexes. The quaternary structure of the extracellular chain
A:Reference number: S19532; MUID:92106333; PMID:1762147
A:Accession: S19618
A:Molecule type: protein
A:Residues: 1-20 <QAB>
A:Experimental source: plume
C:Complex: dodecamers, each consisting of a trimer of tetramers of globin chains; dodecamer
C:Keywords: calcium; dodecamer; heme; homotetramer; oxygen carrier

Query Match      25.2%; Score 29; DB 2; Length 20;
Best Local Similarity 43.8%; Pred. No. 1.2e+03;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY  7  EGAGAEVVKRAAEAKQ 22
    :|:|:|:|:|:|:|:|:|:
Db   1  ELSSSEVKRIDANGKK 16

RESULT 10
PQ0692
photosystem I 18.5K D2 chain - tobacco (Nicotiana tomentosiformis) (fragment)
C:Species: Nicotiana tomentosiformis
C>Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 17-Mar-1999
C:Accession: PQ0692
R:Obokata, J.; Mikami, K.; Hayashida, N.; Nakamura, M.; Sugiyama, M.
Plant Physiol. 102, 1259-1267, 1993
A:Title: Molecular heterogeneity of photosystem I. psaD, psaE, psaF, psaH and psaL are a
A:Reference number: PQ0667; MUID:94105345; PMID:8278548
A:Accession: PQ0692
A:Molecule type: protein
A:Residues: 1-15 <OBO>
C:Superfamily: photosystem I chain II
C:Keywords: chloroplast; photosynthesis; photosystem I; thylakoid

Query Match      24.3%; Score 28; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 7; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY  5  AAEAGAGAEVKRASA 18
    :|:|:|:|:|:|:|:|:|:
Db   1  AEEAATAATKEAEA 14

```

## RESULT 11

```

S38292
30K allergen - rye (fragment)
C:Species: Secale cereale (rye)
C>Date: 19-May-1994 #sequence_revision 27-Feb-1997 #text_change 07-May-1999
C:Accession: S38292
R:Petersen, A.; Schramm, G.; Becker, W.M.; Schlaak, M.
Biol. Chem. Hoppe-Seyler 374, 855-861, 1993
A:Title: Comparison of four grass pollen species concerning their allergens of grass gro
A:Reference number: S38288; MUID:94092339; PMID:7505588
A:Accession: S38292
A:Molecule type: protein
A:Residues: 1-16 <PEI>

```

```

Query Match      24.3%; Score 28; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY  1  GTAPAAEG 8
    :|:|:|:|:|:|:
Db   4  GYAPAAPG 11

```

## RESULT 12

```

S32883
DNA gyrase chain A - Streptomyces spheroides (fragment)
C:Species: Streptomyces spheroides
C>Date: 08-Dec-1993 #sequence_revision 26-May-1995 #text_change 22-Oct-1999
C:Accession: S32883; S29684
R:Thiara, A.S.; Cundliffe, E.
Mol. Microbiol. 8, 495-506, 1993
A:Title: Expression and analysis of two gyrB genes from the novobiocin producer, Streptom
A:Reference number: S32881; MUID:93316846; PMID:8392138
A:Accession: S32883
A:Molecule type: DNA
A:Residues: 1-21 <THI>
A:Cross-references: EMBL:Z17305; NID:g47535; PIDN:CAA78953.1; PID:g47537

```

```

Query Match      24.3%; Score 28; DB 2; Length 21;
Best Local Similarity 71.4%; Pred. No. 1.7e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY  4  PAAEGAG 10
    :|:|:|:|:|:|:
Db   14  PAVEGVG 20

```

## RESULT 13

```

A26376
peptidyl-dipeptidase A (EC 3.4.15.1) - bovine (fragment)
N:Alternate names: ACE; angiotensin-converting enzyme; dipeptidyl carboxypeptidase I; pe
C:Species: Bos primigenius taurus (cattle)
C>Date: 21-May-1988 #sequence_revision 05-Apr-1995 #text_change 30-Jun-1995
C:Accession: A26376; B61477
R:St. Clair, D.K.; Presper, K.A.; Smith, P.L.; Stump, D.C.; Heath, E.C.
Biochem. Biophys. Res. Commun. 141, 968-972, 1986
A:Title: Bovine angiotensin-converting enzyme: amino terminal sequence analysis and preli
A:Reference number: A26376; MUID:87128067; PMID:3028395
A:Accession: A26376
A:Molecule type: protein
A:Residues: 1-22 <STC>
A:Experimental source: lung
R:Bernstein, K.B.; Martin, B.M.; Striker, L.; Striker, G.
Kidney Int. 33, 652-655, 1988
A:Title: Partial protein sequence of mouse and bovine kidney angiotensin converting enzy
A:Reference number: A61477; MUID:88215372; PMID:2835538
A:Accession: B61477
A:Molecule type: protein
A:Residues: 1-22 <BER>
A:Experimental source: lung
C:Superfamily: mammalian peptidyl-dipeptidase A

```

C;Keywords: alternative splicing; blood pressure control; membrane protein; peptidylpe

Query Match 23.0%; Score 26.5; DB 2; Length 22;  
Best Local Similarity 53.8%; Pred. No. 2.9e+03;  
Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 GTAPAAEGAGAEV 13  
||| ||| ||| :  
DB 9 GNFFADE-AGAQI 20

## RESULT 14

PH0918  
T-cell receptor beta chain V-D-J region (isolate 4) - rat (fragment)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 09-Oct-1992 #sequence\_revision 09-Oct-1992 #text\_change 30-May-1997  
C;Accession: PH0918  
R;Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.  
J. Exp. Med. 174, 1467-1476, 1991  
A;Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergic  
A;Reference number: PH0891; MUID:92078857; PMID:1836012  
A;Accession: PH0918  
A;Molecule type: mRNA  
A;Residues: 1-9 <GOL>  
A;Experimental source: concanavalin A-activated lymphoblast  
C;Keywords: T-cell receptor

Query Match 22.6%; Score 26; DB 2; Length 9;  
Best Local Similarity 62.5%; Pred. No. 2.8e+05;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 AAEGAGAE 12  
|: |||  
DB 2 ASSGTGAE 9

## RESULT 15

PS0260  
56K protein 2402 - rice (strain Nihonbare) (fragment)  
C;Species: Oryza sativa (rice)  
C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 11-Apr-1995  
C;Accession: PS0260  
R;Tsugita, A.; Miyatake, N.  
submitted to JIPID, April 1993  
A;Reference number: PS0208  
A;Accession: PS0260  
A;Molecule type: protein  
A;Residues: 1-15 <TSU>  
A;Experimental source: germ  
C;Comment: molecular weight 56K, pI 4.9.

Query Match 22.6%; Score 26; DB 2; Length 15;  
Best Local Similarity 46.7%; Pred. No. 2.4e+03;  
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 6 AEGAGAEVKKASAEA 20  
||| ||| :||  
DB 1 AEAIAAEEXAPAEA 15

Search completed: October 7, 2004, 17:57:26  
Job time : 26 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:36:42 ; Search time 14 seconds  
(without alignments)  
89.263 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAAEGAGAEVKRASAEKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 1550

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	26.1	19	1 ATPB CANFA	P99504 canis famil
2	30	26.1	21	1 PIL3_ECOLI	P13948 escherichia
3	28	24.3	21	1 GYRA STRSH	P50071 streptomyce
4	26.5	23.0	22	1 ACE BOVIN	P12820 bos taurus
5	25	21.7	16	1 KTRC AREMA	P11917 arenicola m
6	25	21.7	20	1 DER6 DERPT	P49277 dermatophag
7	25	21.7	24	1 AMAA_BACTR	P37356 bacillus th
8	24.5	21.3	19	1 FIBA_BISBO	P14441 bison bonas
9	24.5	21.3	19	1 FIBA_CEREL	P14446 cervus elap
10	24.5	21.3	19	1 FIBA_CERNI	P14447 cervus nipp
11	24	20.9	15	1 R13A SPIOL	P82454 spinacia ol
12	24	20.9	15	1 UN04 PINPS	P81673 pinus pinas
13	23	20.0	12	1 H2AX_ONCMY	P83327 oncorhynchu
14	23	20.0	12	1 LMT1_LOCMY	P22395 locusta mig
15	23	20.0	17	1 UP34_UPEMJ	P13066 desulfovibr
16	23	20.0	19	1 PHSL_DESEN	P82041 uperoleia m
17	23	20.0	20	1 ELAS_GADMO	P32197 gadus morbu
18	22	19.1	21	1 TL19_ARATH	P82658 arabidopsis
19	22	19.1	23	1 VG22_BPT2	P21596 bacterioph
20	22	19.1	23	1 VG22_BPT6	P21597 bacterioph
21	21.5	18.7	19	1 FIBA_MUNMU	P14457 muntiacus m
22	21	18.3	8	1 WP1_PERAT	P83195 perkinsus a
23	21	18.3	12	1 PORD_METTM	P80903 methanobact
24	21	18.3	12	1 TA10_TREME	P01371 tremella me
25	21	18.3	14	1 SAP2_ARBPU	P11760 arabacia pun
26	21	18.3	14	1 UC04_WAIZE	P80610 zea mays lm
27	21	18.3	15	1 FIBA_SYNCA	P14463 syncerus ca
28	21	18.3	15	1 RS20_BACST	P59681 bacillus st
29	21	18.3	16	1 CERR_RAT	P23436 rattus norv
30	21	18.3	16	1 FIBA_CERSI	P14535 ceratotheri
31	21	18.3	16	1 FIBA_FELCA	P14450 felis silve
32	21	18.3	16	1 FIBA_HYLLA	P14453 hylobates l
33	21	18.3	16	1 FIBA_NACFU	P12803 macaca fusc

34 21 18.3 16 1 FIBA\_ODOHE  
35 21 18.3 16 1 FIBA\_TAPTE  
36 21 18.3 17 1 FIBA\_PIG  
37 21 18.3 18 1 FIBA\_CAMDR  
38 21 18.3 18 1 FIBA\_LAMGL  
39 21 18.3 18 1 YMDB\_CHLAU  
40 21 18.3 19 1 FIBA\_BUBBU  
41 21 18.3 19 1 FIBA\_SHEEP  
42 21 18.3 20 1 FIBB\_ELEMA  
43 21 18.3 21 1 PEDB\_HYDAT  
44 21 18.3 21 1 PSAL\_SYNPG  
45 21 18.3 21 1 SOD2\_PICAB

P1459 odocoileus  
P14536 tapirus ter  
P14460 sus scrofa  
P14444 camelus dro  
P14454 lama glama  
Q45827 chloroflexu  
P14442 bubalus bub  
P14451 ovis aries  
P14538 elephas max  
P80577 hydra atten  
P31084 synecococc  
P29428 picea abies

## ALIGNMENTS

RESULT 1  
ATPB CANFA  
ID ATPB CANFA STANDARD; PRT; 19 AA.  
AC P99504;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE ATP synthase beta chain, mitochondrial (EC 3.6.3.14) (Fragment).  
GN ATP5B.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE  
RC TISSUE=Heart;  
RX MEDLINE=98163340; PubMed=9504812;  
RA Dunn M.J., Corbett J.M., Wheeler C.H.;  
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of  
dog heart proteins.";  
RL Electrophoresis 18:2795-2802(1997).  
CC -!- FUNCTION: Produces ATP from ADP in the presence of a proton  
gradient across the membrane. The beta chain is the catalytic  
subunit.  
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +  
H(+) (Out).  
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic  
core - and CF(0) - the membrane proton channel. CF(1) has five  
subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)  
has three main subunits: a, b and c.  
CC -!- SUBCELLULAR LOCATION: Mitochondrial.  
CC -!- SIMILARITY: Belongs to the ATPase alpha/beta chains family.  
CC HSC-2DPAGE; P99504; DOG.  
DR InterPro: IPR000194; ATPase\_a/bcentre.  
DR PROSITE; PS00152; ATPASE\_ALPHA\_BETA; PARTIAL.  
KW ATP synthesis; CF(1); Hydrogen ion transport; Hydrolase; ATP-binding;  
Mitochondrion.  
FT UNSURE 8 8  
FT NON\_TER 17 19  
FT NON\_TER 19 19  
SQ SEQUENCE 19 AA; 1871 MW; BB9C163FDC60BB42 CRC64;  
Query Match 26.1%; Score 30; DB 1; Length 19;  
Best Local Similarity 42.9%; Pred. No. 4.2e+02;  
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
QY 2 TAPAAEGAGAEVKR 15  
|:::|  
Db 4 TSPSPKGAAXXR 17  
RESULT 2  
PIL3\_ECOLI  
ID PIL3\_ECOLI STANDARD; PRT; 21 AA.  
AC P13948;  
DT 01-JAN-1990 (Rel. 13, Created)

```
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1999 (Rel. 35, Last annotation update)
DE Fimbrial protein precursor (Pilin) (Fragment).
GN TRAA.
OS Escherichia coli.
OG Plasmid ColB4-K98.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87008371; PubMed=3531163;
RA Finlay B.B., Frost L.S., Parachych W.;
RT "Origin of transfer of IncF plasmids and nucleotide sequences of the
RT type II oriT, traM, and traY alleles from ColB4-K98 and the type IV
RT traY allele from R100-1.";
RL J. Bacteriol. 168:132-139(1986).
CC -!- FUNCTION: PROPILIN IS THE PRECURSOR OF THE SEXPIILUS SUBUNIT. THE
CC SEXPIILUS ARE FILAMENTOUS SURFACE APPENDAGES REQUIRED FOR CELL-TO-
CC CELL CONTACT DURING BACTERIAL CONJUGATION.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: M15135; AAB04666.1; -.
DR InterPro; IPR008873; TrAa.
DR Pfam; PF05513; TrAa; 1.
KW Plasmid; Conjugation; Fimbria.
FT PROPEP 1 >21
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2198 MW; 076CDD0C5E9D14EA CRC64;
Query Match 26.1%; Score 30; DB 1; Length 21;
Best Local Similarity 46.2%; Pred. No. 4.6e+02;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 5 AAEAGAGAEVKRAS 17
Db :||| |||
6 SVQGASAPVKKS 18
RESULT 3
GYRA_STRSH STANDARD; PRT; 21 AA.
AC P50071;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE DNA gyrase subunit A (EC 5.99.1.3) (Fragment).
GN GYRA.
OS Streptomyces sphearoides.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=195949;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NCIB 11891;
RX MEDLINE=93316846; PubMed=8392138;
RA Thiara A.S., Cundliffe E.;
RT "Expression and analysis of two gyrB genes from the novobiocin
RT producer, Streptomyces sphearoides.";
RL Mol. Microbiol. 8:495-506(1993).
CC -!- FUNCTION: DNA gyrase negatively supercoils closed circular double-
CC stranded DNA in an ATP-dependent manner and also catalyzes the
CC interconversion of other topological isomers of double-stranded
CC DNA rings, including catenanes and knotted rings.
CC -!- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
```

```
CC of double-stranded DNA.
CC -!- SUBUNIT: Made up of two chains. The A chain is responsible for DNA
CC breakage and rejoining; the B chain catalyzes ATP hydrolysis. The
CC enzyme forms an A2B2 tetramer.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: Z17305; CAA78953.1; -.
DR PIR; S32883; S32883.
KW Topoisomerase; Isomerase; DNA-binding.
RN NON_TER 21 21
SQ SEQUENCE 21 AA; 2203 MW; 5E2F9DEA8DB83697 CRC64;
Query Match 24.3%; Score 28; DB 1; Length 21;
Best Local Similarity 71.4%; Pred. No. 8.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4 PAAGGAG 10
Db |||||
14 PAVEGVG 20
RESULT 4
ACE_BOVIN STANDARD; PRT; 22 AA.
ID ACE_BOVIN
AC P12820;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Angiotensin-converting enzyme, somatic (EC 3.4.15.1) (ACE) (Dipeptidyl
DE carboxypeptidase I) (Kininase II) (Fragment).
GN ACE OR DCPI.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Lung;
RX MEDLINE=87128067; PubMed=3028395;
RA St Clair D.K., Presper K.A., Smith P.L., Stump D.C., Heath E.C.;
RT "Bovine angiotensin-converting enzyme: amino-terminal sequence
RT analysis and preliminary characterization of a hybridization-selected
RT primary translation product.";
RL Biochem. Biophys. Res. Commun. 141:968-972(1986).
CC -!- FUNCTION: Converts angiotensin I to angiotensin II by release of the
CC the terminal His-Ieu, this results in an increase of the
CC vasoconstrictor activity of angiotensin.
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal dipeptide,
CC oligopeptide-|-Xaa-Xbb, when Xaa is not Pro, and Xbb is neither
CC Asp nor Glu. Converts angiotensin I to angiotensin II.
CC -!- COFACTOR: Binds 2 zinc ions per subunit (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family M2.
DR PIR; A26376; A26376.
DR MEROPS; M02.001; -.
DR InterPro; IPR006025; Pept M Zn_Bs.
DR PROSITE; PS00142; ZINC_PROTEASE; PARTIAL.
KW Hydrolase; Metalloprotease; Carboxypeptidase; Zinc.
FT NON_TER 22 22
SQ SEQUENCE 22 AA; 2271 MW; 77B0AC6A91A3893F CRC64;
Query Match 23.0%; Score 26.5; DB 1; Length 22;
Best Local Similarity 53.8%; Pred. No. 1.4e+03;
Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;
QY 1 GTAPAAEGAGAEV 13
```





```

RT "Studies on fibrinopeptides from mammals.";
RL Acta Chem. Scand. 19:1789-1791(1965).
CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC which cleaves fibrinopeptides A and B from alpha and beta chains,
CC and thus exposes the N-terminal polymerization sites responsible
CC for the formation of the soft clot.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 19 FIBRINOPEPTIDE A.
FT NON TER 19 19
SQ SEQUENCE 19 AA; 1836 MW; 9BA55A0F473B59C5 CRC64;

Query Match 21.3%; Score 24.5; DB 1; Length 19;
Best Local Similarity 46.7%; Pred. No. 2.3e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

Qy 1 GTAPD-----AEGAG 10
|:|:|:|:|
Db 3 GSDPASGDFLAEGGG 17

RESULT 9
FIBA CEREL
ID FIBA CEREL STANDARD; PRT; 19 AA.
AC P1446;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN FGA.
OS Cervus elaphus (Red deer), and
OS Cervus elaphus nelsoni (American elk).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
OC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=9860, 9864;
RN [1]
RP SEQUENCE.
RC SPECIES=C.elaphus;
RA Moss G.A., Doolittle R.F.;
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";
RL Arch. Biochem. Biophys. 122:674-684(1967).
EN [2]
RP SEQUENCE.
RC SPECIES=C.e.nelsoni;
RA Moss G.A., Doolittle R.F.;
RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";
RL Arch. Biochem. Biophys. 122:674-684(1967).
EN [2]
RP SEQUENCE.
CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC which cleaves fibrinopeptides A and B from alpha and beta chains,
CC and thus exposes the N-terminal polymerization sites responsible
CC for the formation of the soft clot.
KW Blood coagulation; Plasma.
FT PEPTIDE 1 19 FIBRINOPEPTIDE A.
FT NON TER 19 19
SQ SEQUENCE 19 AA; 1808 MW; 9BA54C26873B59C5 CRC64;

Query Match 21.3%; Score 24.5; DB 1; Length 19;
Best Local Similarity 46.7%; Pred. No. 2.3e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

Qy 1 GTAPD-----AEGAG 10
|:|:|:|:|
Db 3 GSDPASGDFLAEGGG 17

```

```

RESULT 10
FIBA CERNI
ID FIBA CERNI STANDARD; PRT; 19 AA.
AC P1447;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha chain [Contains: Fibrinopeptide A] (Fragment).
GN FGA.
OS Cervus nippon (Sika deer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;
OC Cervidae; Cervinae; Cervus.
OX NCBI_TaxID=9863;
RN [1]
RP SEQUENCE.
RA Blomback B., Blomback M., Grondahl N.J., Holmberg E.;
RT "Structure of fibrinopeptides-its relation to enzyme specificity and
RT phylogeny and classification of species.";
RL Ark. Kemi 25:411-428(1966).
CC -!- FUNCTION: Fibrinogen has a double function: yielding monomers that
CC polymerize into fibrin and acting as a cofactor in platelet
CC aggregation.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: Conversion of fibrinogen to fibrin is triggered by thrombin,
CC which cleaves fibrinopeptides A and B from alpha and beta chains,
CC and thus exposes the N-terminal polymerization sites responsible
CC for the formation of the soft clot.
KW Blood coagulation; Plasma.
FT NON TER 19 19
SQ SEQUENCE 19 AA; 1822 MW; 9BA40926873B59C5 CRC64;

Query Match 21.3%; Score 24.5; DB 1; Length 19;
Best Local Similarity 46.7%; Pred. No. 2.3e+03;
Matches 7; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

Qy 1 GTAPD-----AEGAG 10
|:|:|:|:|
Db 3 GSDPASSEFLAEGGG 17

RESULT 11
R13A SPIOL
ID R13A SPIOL STANDARD; PRT; 15 AA.
AC P82454;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 60S ribosomal protein L13a (Fragment).
GN RPL13A.
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE.
RC STEAIN=cv. Alvaro; TISSUE=Leaf;
RA Yamaguchi K., Subramanian A.R.;
RT "N-terminal sequence of spinach cytosolic 60S ribosomal protein
RT L13a.";
RL Submitted (APR-2000) to Swiss-Prot.
CC -!- SIMILARITY: Belongs to the L13P family of ribosomal proteins.
DR InterPro; IPR005822; Ribosomal_L13.
DR PROSITE; PS00783; RIBOSOMAL_L13; PARTIAL.
KW Ribosomal protein.
FT NON TER 15 15
SQ SEQUENCE 15 AA; 1489 MW; C7B9C80F5A099EB3 CRC64;

Query Match 20.9%; Score 24; DB 1; Length 15;

```



RN [1]  
 RP SEQUENCE, AND MASS SPECTROMETRY.  
 RA TISSUE=Skin secretion; Bowle J.H., Tyler M.J., Wallace J.C.;  
 RT "New antibiotic uperin peptides from the dorsal glands of the  
 RT Australian toadlet Uperoleia mjobergii.",  
 RL Aust. J. Chem. 49:1325-1331(1996).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.  
 CC -!- MASS SPECTROMETRY: MW=1735; METHOD=FAB.  
 KW Amphibian defense peptide; Amidation.  
 FT MOD RES 17 17 AMIDATION.  
 SQ SEQUENCE 17 AA; 1737 MW; 6F61E4834375DE1B CRC64;  
  
 Query Match 20.0%; Score 23; DB 1; Length 17;  
 Best Local Similarity 35.7%; Pred. No. 3.4e+03;  
 Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
  
 QY 8 GAGAEVKRASAFAK 21  
 | | : : | |  
 Db 1 GVGDLIRKAVAAIK 14

Search completed: October 7, 2004, 17:54:03  
 Job time : 15 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:37:47 ; Search time 72 Seconds  
(without alignments)  
105.173 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAEAGAGAEVKRASAEAKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 10229

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTRMBL 25:\*

2: sp\_archaea:\*

3: sp\_bacteria:\*

4: sp\_fungi:\*

5: sp\_human:\*

6: sp\_invertebrate:\*

7: sp\_mammal:\*

8: sp\_mmc:\*

9: sp\_organelle:\*

10: sp\_phage:\*

11: sp\_plant:\*

12: sp\_rodent:\*

13: sp\_virus:\*

14: sp\_vertebrate:\*

15: sp\_unclassified:\*

16: sp\_rvirus:\*

17: sp\_bacteriap:\*

18: sp\_archaea:\*

19: sp\_bacteria:\*

20: sp\_fungi:\*

21: sp\_human:\*

22: sp\_invertebrate:\*

23: sp\_mammal:\*

24: sp\_mmc:\*

25: sp\_organelle:\*

26: sp\_phage:\*

27: sp\_plant:\*

28: sp\_rodent:\*

29: sp\_virus:\*

30: sp\_vertebrate:\*

31: sp\_unclassified:\*

32: sp\_rvirus:\*

33: sp\_bacteriap:\*

34: sp\_archaea:\*

35: sp\_bacteria:\*

36: sp\_fungi:\*

37: sp\_human:\*

38: sp\_invertebrate:\*

39: sp\_mammal:\*

40: sp\_mmc:\*

41: sp\_organelle:\*

42: sp\_phage:\*

43: sp\_plant:\*

44: sp\_rodent:\*

45: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	37	32.2	24	10 Q9S8S0	Q9S8S0 chlamydomon
2	36	31.3	22	5 Q9TWU8	Q9TWU8 crithidia f
3	32	27.8	23	4 P78489	P78489 homo sapien
4	32	27.8	23	4 Q9UCN3	Q9UCN3 homo sapien
5	31	27.0	20	4 Q15988	Q15988 homo sapien
6	31	27.0	23	4 Q43519	Q43519 homo sapien
7	31	27.0	24	2 P81151	P81151 desulfovibr
8	30.5	26.5	20	10 Q9S8X9	Q9S8X9 phleum prat
9	30	26.1	15	10 Q9SQ16	Q9SQ16 oryza sativ
10	30	26.1	17	11 Q8K407	Q8K407 rattus norv
11	30	26.1	24	12 Q83877	Q83877 norwalk vir
12	29	25.2	18	11 Q55167	Q55167 rattus norv
13	29	25.2	22	8 Q9T217	Q9T217 nicotiana s
14	29	25.2	22	10 Q48861	Q48861 oryza sativ
15	28	24.3	15	8 Q9T210	Q9T210 nicotiana t
16	28	24.3	16	11 Q9QUY8	Q9QUY8 cricetulus

17	28	24.3	18	10 Q9S8Y9	Q9S8Y9 poa pratens
18	27	23.5	12	2 Q9X6Y0	Q9X6Y0 aquifex pyr
19	27	23.5	21	6 Q9TR11	Q9TR11 bos taurus
20	27	23.5	21	10 Q9S7S6	Q9S7S6 triticum ae
21	27	23.5	21	10 Q42417	Q42417 triticum ae
22	27	23.5	21	10 Q42501	Q42501 triticum ae
23	27	23.5	22	11 Q9QWS7	Q9QWS7 mus musculu
24	27	23.5	22	13 Q9PS65	Q9PS65 oncorhynch
25	26.5	23.0	21	12 Q10422	Q10422 influenza a
26	26.5	23.0	24	4 Q9UNM2	Q9UNM2 homo sapien
27	26	22.6	15	2 Q9R5T1	Q9R5T1 flavobacter
28	26	22.6	18	10 Q40499	Q40499 nicotiana t
29	26	22.6	20	8 Q9T218	Q9T218 nicotiana s
30	26	22.6	20	10 Q9S8B7	Q9S8B7 oryza sativ
31	26	22.6	20	10 Q9S7R1	Q9S7R1 oryza glabe
32	26	22.6	21	10 Q41566	Q41566 triticum ae
33	26	22.6	23	12 Q84331	Q84331 simian viru
34	25	21.7	16	8 Q9T216	Q9T216 nicotiana s
35	25	21.7	16	10 P82453	P82453 spinacia ol
36	25	21.7	16	13 Q9PRU6	Q9PRU6 gallus gall
37	25	21.7	17	13 Q9PRU7	Q9PRU7 gallus gall
38	25	21.7	19	2 Q9RSH4	Q9RSH4 rhodospiril
39	25	21.7	20	4 Q9UMI8	Q9UMI8 homo sapien
40	25	21.7	21	4 Q86UF6	Q86UF6 homo sapien
41	25	21.7	21	10 Q41559	Q41559 triticum ae
42	24.5	21.3	19	2 Q06140	Q06140 neisseria m
43	24	20.9	13	2 Q7X205	Q7X205 alteromonas
44	24	20.9	15	5 P82211	P82211 bombyx mori
45	24	20.9	15	11 Q9QV72	Q9QV72 mus sp. int

## ALIGNMENTS

## RESULT 1

Q9S8S0	PRELIMINARY;	PRT;	24 AA.
AC	Q9S8S0		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)		
DE	Hydrogenase (EC 1.12.-.-) (Fragment).		
OS	Chlamydomonas reinhardtii.		
OC	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;		
OC	Chlamydomonadaceae; Chlamydomonas.		
OX	NCBI_TaxID=3055;		
RN	[1]		
RP	SEQUENCE.		
RX	MEDLINE=93292509; PubMed=8513797;		
RA	Happe T., Naber J.D.;		
RT	"Isolation, characterization and N-terminal amino acid sequence of		
RT	hydrogenase from the green alga Chlamydomonas reinhardtii.";		
RL	Bur. J. Biochem. 214:475-481(1993).		
DR	Pir, S33302; S33302.		
SQ	SEQUENCE 24 AA; 2427 MW; A582111FEALB69EC CRC64;		

Query Match 32.2%; Score 37; DB 10; Length 24;  
Best Local Similarity 52.9%; Pred. No. 3.5e+02;

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY	3	APAAEGAGAEVKRASAE 19
DB	2	APAAEAPLSHVQQAALAE 18

## RESULT 2

Q9TWU8	PRELIMINARY;	PRT;	22 AA.
ID	Q9TWU8		
AC	Q9TWU8		
DT	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)		
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)		
DE	Kinetoplast DNA-associated protein (Fragment).		

OS Crithidia fasciculata.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Crithidia.  
OX NCBI\_TaxID=5656;  
RN [1]  
RN SEQUENCE.  
RP MEDLINE=93189582; PubMed=8446592;  
RX Xu C., Ray D.S.;  
RA "Isolation of proteins associated with kinetoplast DNA networks in  
RT vivo";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:1786-1789(1993).  
DR PIR; D47256;  
SQ SEQUENCE 22 AA; 1991 MW; 5B1C4941E08F9B2 CRC64;  
  
Query Match 31.3%; Score 36; DB 5; Length 22;  
Best Local Similarity 42.1%; Pred. No. 4.4e+02;  
Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;  
  
QY 3 APAEAGAGAEVVKRASAIAK 21  
||| : : : :  
Db 4 APAKKAAPKAASAPAK 22  
||| : : : :  
  
RESULT 3  
P78489 PRELIMINARY; PRT; 23 AA.  
AC P78489;  
DT 01-MAY-1997 (TReMBLrel. 03, Created)  
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE SP-A1 epsilon (Fragment).  
GN SP-A1 OR SP-A2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=94234365; PubMed=8179012;  
RA McCormick S.M., Boggs V., Mendelson C.R.;  
RT "Characterization of mRNA transcripts and organization of human SP-A1  
RT and SP-A2 genes";  
RL Am. J. Physiol. 266:L354-L366(1994).  
DR EMBL; S69690; AAB30738.1; -  
DR EMBL; S69680; AAB30730.1; -  
DR EMBL; S69681; AAB30731.1; -  
DR EMBL; S69685; AAB30734.1; -  
DR EMBL; S69686; AAB30735.1; -  
DR PIR; I64839; I64839.  
FT NON\_TER 23 23  
SQ SEQUENCE 23 AA; 2419 MW; 8C4FE026553AD2C CRC64;  
  
Query Match 27.8%; Score 32; DB 4; Length 23;  
Best Local Similarity 70.0%; Pred. No. 1.5e+03;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 5 AAEGAGAEVK 14  
||| |||  
Db 14 AASGAEVK 23  
||| |||  
  
RESULT 4  
Q9UCN3 PRELIMINARY; PRT; 23 AA.  
AC Q9UCN3;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Interleukin-1 beta converting enzyme (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]

RP SEQUENCE.  
RX MEDLINE=92337439; PubMed=1321594;  
RA Kronheim S.R., Muma A., Greenstreet T., Glackin P.J., Van Ness K.,  
RA March C.J., Black R.A.;  
RT "Purification of interleukin-1 beta converting enzyme, the protease  
RT that cleaves the interleukin-1 beta precursor";  
RL Arch. Biochem. Biophys. 296:698-703(1992).  
DR GO; GO:0003824; F:catalytic activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR000873; AMP-bind.  
DR PROSITE; PS00455; AMP\_BINDING; 1.  
SQ SEQUENCE 23 AA; 2322 MW; 63C2E59F1F3102D6 CRC64;  
  
Query Match 27.8%; Score 32; DB 4; Length 23;  
Best Local Similarity 36.8%; Pred. No. 1.6e+03;  
Matches 7; Conservative 2; Mismatches 10; Indels 0; Gaps 0;  
  
QY 4 PAEAGAGAEVVKRASAIAK 22  
||| : : : :  
Db 5 PTSGSGNVKLXLEXAQ 23  
||| : : : :  
  
RESULT 5  
Q15988 PRELIMINARY; PRT; 20 AA.  
AC Q15988;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)  
DE HLX1 protein (fragment).  
GN HLX1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=93194183; PubMed=8095486;  
RA Nishimura D.Y., Purchio A.F., Murray J.C.;  
RT "Linkage localization of TGFB2 and the human homeobox gene HLX1 to  
RT chromosome 1q";  
RL Genomics 15:357-364(1993).  
DR EMBL; S56767; AADI3883.1; -  
FT NON\_TER 1 1  
SQ SEQUENCE 20 AA; 2012 MW; 6BB655F09B5B5AE4 CRC64;  
  
Query Match 27.0%; Score 31; DB 4; Length 20;  
Best Local Similarity 75.0%; Pred. No. 1.9e+03;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GTAPAAEG 8  
||| |||  
Db 4 GGAPADG 11  
||| |||  
  
RESULT 6  
O43519 PRELIMINARY; PRT; 23 AA.  
AC O43519;  
DT 01-JUN-1998 (TReMBLrel. 06, Created)  
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE RET proto-oncogene (fragment).  
GN RET.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=98086132; PubMed=9426223;  
RA Patrone G., Puliti A., Bocciaardi R., Ravazzolo R., Romeo G.;  
RT "Sequence and characterisation of the RET proto-oncogene 5' flanking

RT region: analysis of retinoic acid responsiveness at the  
 RL transcriptional level.";  
 RN FEBS Lett. 419:76-82(1997).  
 RP [2]  
 RX MEDLINE=99054655; PubMed=9840920;  
 RA Munnes M., Patrone G., Schmitz B., Romeo G., Doerfler W.;  
 RT "A 5'-CG-3'-rich region in the promoter of the transcriptionally  
 frequently silenced RET protooncogene lacks methylated cytidine  
 residues.";  
 RL Oncogene 17:2573-2583(1998).  
 DR EMBL AF032124; AAB97168.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 23 23  
 SQ SEQUENCE 23 AA; 2129 MW; 1DBAF1BF80E4C66D CRC64;

Query Match 27.0%; Score 31; DB 4; Length 23;  
 Best Local Similarity 46.7%; Pred. No. 2.2e+03;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 8 GAGAEVKRASAEAKQ 22  
 |||:|:|:|:|:  
 Db 9 GAASAVAAAAARQ 23

RESULT 7  
 P81151  
 ID P81151 PRELIMINARY; PRT; 24 AA.  
 AC P81151;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Cytochrome C3 (Fragment).  
 OS Desulfovibrio vulgaris.  
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacteriales;  
 OC Desulfovibrionaceae; Desulfovibrio.  
 OX NCBI\_TaxID=881;  
 RN [1]  
 RP SEQUENCE.

RX MEDLINE=93272123; PubMed=8388770;  
 RA Kwok D.Y., Vedwick T.S., McCue A.F., Gevertz D.;  
 RT "Rapid comparison of the cytochrome c3 gene from nine strains of  
 Desulfovibrio vulgaris using polymerase chain reaction  
 amplification.";  
 RL Can. J. Microbiol. 39:402-411(1993).  
 CC -!- FUNCTION: PARTICIPATES IN SULFATE RESPIRATION COUPLED WITH  
 PHOSPHORYLATION BY TRANSFERRING ELECTRONS FROM THE ENZYME  
 DEHYDROGENASE TO FERREDOXIN.  
 CC -!- PTM: BINDS FOUR NONPARALLEL HEME GROUPS PER MOLECULE.  
 CC -!- SIMILARITY: TO OTHER C3-TYPE CYTOCHROMES.  
 DR GO; GO:0009061; P:anaerobic respiration; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 KW Electron transport; Sulfate respiration; Heme.  
 FT NON\_TER 1 1  
 FT NON\_TER 24 24  
 SQ SEQUENCE 24 AA; 2355 MW; B4F438AFF102472 CRC64;

Query Match 27.0%; Score 31; DB 2; Length 24;  
 Best Local Similarity 42.9%; Pred. No. 2.3e+03;  
 Matches 9; Conservative 3; Mismatches 5; Indels 4; Gaps 1;

QY 1 GTAPAAEGAGAEVKRASAEAK 21  
 |||:|:|:|:|:  
 Db 1 GNAPAAD----MVLKAPGDAK 17

RESULT 8  
 Q9S8X9  
 ID Q9S8X9 PRELIMINARY; PRT; 20 AA.  
 AC Q9S8X9;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Allergen PHLP V (Fragment).  
 OS Phleum pratense (Common timothy).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;  
 OC Aveneae; Phleum.  
 OX NCBI\_TaxID=15957;  
 RN [1]  
 RP SEQUENCE  
 RX MEDLINE=92353723; PubMed=1643437;  
 RA Petersen A., Becker W.M., Schlaak M.;  
 RT "Characterization of isoforms of the major allergen Phl p V by two-  
 dimensional immunoblotting and microsequencing.";  
 RL Int. Arch. Allergy Immunol. 98:105-109(1992).  
 DR PIR: S38294; S38294.  
 SQ SEQUENCE 20 AA; 1731 MW; F0DD5E9B482D9836 CRC64;

Query Match 26.5%; Score 30.5; DB 10; Length 20;  
 Best Local Similarity 52.9%; Pred. No. 2.3e+03;  
 Matches 9; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

QY 1 GTAPAA-EGAGAEVKRA 16  
 |||:|:|:|:|:  
 Db 4 GYAPATPAAGAGAEAGKA 20

RESULT 9  
 Q9SQI6  
 ID Q9SQI6 PRELIMINARY; PRT; 15 AA.  
 AC Q9SQI6;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Translation elongation factor (fragment).  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Li Z.-Y., Chen S.-Y.;  
 RT "Inducible expression of translation elongation factor 1A gene in rice  
 seedlings in response to environmental stresses.";  
 RL Acta Bot. Sin. 41:800-806(1999).  
 DR EMBL AF067195; AAC79991.1; -.  
 DR Gramene; Q9SQI6; -.  
 DR GO; GO:0003746; F:translation elongation factor activity; IEA.  
 KW Elongation factor.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 15 AA; 1514 MW; 95E5C46F069C9775 CRC64;

Query Match 26.1%; Score 30; DB 10; Length 15;  
 Best Local Similarity 50.0%; Pred. No. 2e+03;  
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 10 GAEVKRSAGEAK 21  
 |||:|:|:|:|:  
 Db 4 GAKVTQAAKXK 15

RESULT 10  
 Q8K407  
 ID Q8K407 PRELIMINARY; PRT; 17 AA.  
 AC Q8K407;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Tissue-specific calpain 2 (fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley;  
RX MEDLINE=22145722; PubMed=12150941;  
RA Duan W.R., Ito M., Lee E.J., Chien P.-Y., Jameson J.L.;  
RT "Estrogen regulates a tissue-specific calpain in the anterior  
RT pituitary";  
RL Biochem. Biophys. Res. Commun. 295:261-266(2002).  
DR EMBL; AF514419; AAM94284.1; -.  
FT NON\_TER 17  
SQ SEQUENCE 17 AA; 1630 MW; 2027F409989F3A96 CRC64;

Query Match 26.1%; Score 30; DB 11; Length 17;  
Best Local Similarity 53.3%; Pred. No. 2.2e+03;  
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 AAGAGAEVVKASAE 19  
Db 2 AALAGVSKQRAVAE 16  
|||:|||||

RESULT 11  
Q83877 PRELIMINARY; PRT; 24 AA.  
AC Q83877;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
DE Putative ORF3 (Fragment).  
OS Norwalk virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=11983;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SRV-KY-89/89/J;  
RX MEDLINE=94335115; PubMed=8057474;  
RA Wang J., Jiang X., Madore H.P., Gray J., Desselberger U., Ando T.,  
RA Seto Y., Oishi I., Lew J.F., Green K.Y., et al.;  
RT "Sequence diversity of small, round-structured viruses in the Norwalk  
RT virus group";  
RL J. Virol. 68:5982-5990(1994).  
DR EMBL; L23828; AAA59230.1; -.  
DR InterPro: IPR004278; RNA capsid.  
DR Pfam: PF03035; RNA\_capsid; 1.  
FT NON\_TER 24  
SQ SEQUENCE 24 AA; 2142 MW; 3F47CEE864CA957B CRC64;

Query Match 26.1%; Score 30; DB 12; Length 24;  
Best Local Similarity 58.3%; Pred. No. 3.2e+03;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 TAPAAEGAGAEV 13  
Db 13 TAGSALGAGIQV 24  
|||:|||||

RESULT 12  
O55167 PRELIMINARY; PRT; 18 AA.  
AC O55167;  
DT 01-JUN-1998 (TRENBLrel. 06, Created)  
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)  
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)  
DE ARE1 (Fragment).  
GN ARE1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LEW.1W/Gun;  
RX MEDLINE=99009314; PubMed=9790748;

Walter L., Guenther E.;  
"Identification of a novel highly conserved gene in the centromeric  
part of the major histocompatibility complex.";  
Genomics 52:298-304(1998).  
EMBL; AJ223831; CAAL1568.1; -.  
FT NON\_TER 18  
SQ SEQUENCE 18 AA; 1505 MW; B88156A2048C3388 CRC64;

Query Match 25.2%; Score 29; DB 11; Length 18;  
Best Local Similarity 75.0%; Pred. No. 3.2e+03;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 APAAEGAG 10  
Db 8 AAAAQGAG 15  
|||:|||||

RESULT 13  
Q9T217 PRELIMINARY; PRT; 22 AA.  
AC Q9T217;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)  
DE 14.3 photosystem I PSAE protein (Fragment).  
OS Nicotiana sylvestris (Wood tobacco).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Nicotiana.  
OX NCBI\_TaxID=4096;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=94105345; PubMed=8278548;  
RA Obokata J., Mikami K., Hayashida N., Nakamura M., Sugiura M.;  
RT "Molecular heterogeneity of photosystem I. psad, psae, psaf, psah, and  
RT psal are all present in isoforms in Nicotiana spp.";  
RL Plant Physiol. 102:1259-1267(1993).  
SQ SEQUENCE 22 AA; 2033 MW; 488E17845223FFF7 CRC64;

Query Match 25.2%; Score 29; DB 8; Length 22;  
Best Local Similarity 44.4%; Pred. No. 4e+03;  
Matches 8; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 3 APAAEGAGAEVVKASAE 20  
Db 5 APFAAATAEPAEAPVKA 22  
|||:|||||

RESULT 14  
O48861 PRELIMINARY; PRT; 22 AA.  
AC O48861;  
DT 01-JUN-1998 (TRENBLrel. 06, Created)  
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)  
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)  
DE Homeobox protein (Fragment).  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IRBB21;  
RA Liu G., Yang J., Zhai W., He P., Li X., Lu J., Li S., Zhu L.;  
RT "Mapping study of rice homeobox gene family";  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF003600; AAB94485.1; -.  
DR Gramene; O48861; -.  
FT NON\_TER 1  
FT NON\_TER 22  
SQ SEQUENCE 22 AA; 2337 MW; B18E95E5952F6CE9C CRC64;



Query Match 25.2%; Score 29; DB 10; Length 22;  
Best Local Similarity 40.0%; Pred. No. 4e+03;  
Matches 6; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 GTAPAAEGAGAEVKR 15  
| | | | | : | :  
DB 1 GKAAAEESTGLDAKQ 15

## RESULT 15

Q9T2I0 PRELIMINARY; PRT; 15 AA.  
AC Q9T2I0;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE 18.5 kDa photosystem I PSAD protein (Fragment).  
OS Nicotiana tomentosiformis (Tobacco).  
OG Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC Lamiales; Solanales; Solanaceae; Nicotiana.  
OX NCBI\_TaxID=4098;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=94105345; PubMed=8278548;  
RA Obokata J., Mikami K., Hayashida N., Nakamura M., Sugiura M.;  
RT "Molecular heterogeneity of photosystem I: psaD, psaE, psaF, and  
RT psaL are all present in isoforms in Nicotiana spp.";  
RL Plant Physiol. 102:1259-1267 (1993).  
DR PIR; PQ0692; PQ0692.  
SQ SEQUENCE 15 AA; 1429 MW; D2388E480B5760A8 CRC64;

Query Match 24.3%; Score 28; DB 8; Length 15;  
Best Local Similarity 50.0%; Pred. No. 3.7e+03;  
Matches 7; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 AAEGAGAEVKRASA 18  
| | | | | : | :  
DB 1 AEEAAAATKEAEA 14

Search completed: October 7, 2004, 17:56:37  
Job time : 76 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:36:12 ; Search time 84.5 Seconds  
(without alignments)  
80.250 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115  
Sequence: 1 GTAPAAEGAGAEVKSASAKQAF 24

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 538975

Minimum DB seq length: 0  
Maximum DB seq length: 24

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	115	100.0	24	6	ABP97766 Scrambled
2	44	38.3	18	7	ADD23833 Breast ca
3	38	33.0	24	5	ABP62492 Human imm
4	36	31.3	8	4	ABP12991 HIV A02 s
5	36	31.3	8	4	ABP20973 HIV A03 m
6	36	31.3	8	7	ADD56874 HLA bindi
7	36	31.3	9	4	ABP16782 HIV B07 s
8	36	31.3	9	4	ABP17983 HIV B58 s
9	36	31.3	10	7	ABP13017 HIV A02 s
10	36	31.3	11	4	ABP13067 HLA bindi
11	36	31.3	11	7	ADD57290 HIV A02 s
12	36	31.3	20	2	AAR43045 N-termina
13	36	31.3	22	5	AAR95258 Vector en
14	36	31.3	22	5	AAU99665 pET15b en
15	36	31.3	22	6	ABU09796 Fusion pr
16	36	31.3	22	6	ABG72906 Amino-ter
17	36	31.3	23	2	AAR41278 pET-3A de
18	36	31.3	15	4	AAG62627 Human RNA
19	35	30.4	24	5	ABP62462 Human imm
20	35	30.4	24	5	ABP62465 Human imm
21	35	30.4	24	5	ABP62463 Human imm
22	35	30.4	24	5	ABP62481 Human imm
23	35	30.4	24	5	ABP62464 Human imm
24	35	30.4	24	5	ABP62491 Human imm
25	35	30.4	24	5	ABP62491 Human imm

26	35	30.4	24	5	ABP62480 Human imm
27	35	30.4	24	5	ABP62479 Human imm
28	35	30.4	24	5	ABP62488 Human imm
29	35	30.4	24	5	ABP62490 Human imm
30	34	29.6	14	6	ABP74859 Proteome
31	34	29.6	20	4	ABB76936 Rat VG1-1
32	34	29.6	21	2	AAV20200 Human bet
33	34	29.6	23	6	AAE38088 Human COU
34	34	29.6	24	5	ABP62478 Human imm
35	33.5	29.1	10	4	AAG95102 Human com
36	33.5	29.1	10	4	AAG95102 Human com
37	33.5	29.1	24	5	ABG65791 Plant pot
38	33	28.7	8	4	ABP12990 HIV A02 s
39	33	28.7	8	4	ABP20942 HIV A02 s
40	33	28.7	9	4	ABP16781 HIV B07 s
41	33	28.7	9	4	ABP17977 HIV B58 s
42	33	28.7	9	4	ABP13016 HIV A02 s
43	33	28.7	10	4	ABP18930 HIV B62 s
44	33	28.7	10	4	ABP16789 HIV B07 s
45	33	28.7	10	7	ADD57307 HLA bindi

## ALIGNMENTS

RESULT 1  
ABP97766  
ID ABP97766 standard; peptide; 24 AA.  
XX  
AC ABP97766;  
XX  
DT 11-AUG-2003 (first entry)  
XX  
DE Scrambled myristoylated N-terminal sequence (MANS) peptide.  
XX  
KW Myristoylated N-terminal sequence; MANS; MARCKS; mucus secretion;  
KW inflammatory mediator; inflammation; respiratory disease; asthma;  
KW chronic bronchitis; chronic obstructive pulmonary disease; COPD;  
KW bowel disease; irritable bowel syndrome; Crohn's disease;  
KW ulcerative colitis; skin disease; rosacea; eczema; psoriasis; acne;  
KW autoimmune disease; pain; arthritis; cystic fibrosis.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1 /note= "N-terminal myristate chain attached"  
XX  
WO2003000027-A2.  
XX  
03-JAN-2003.  
XX  
26-JUN-2002; 2002WO-US022270.  
XX  
26-JUN-2001; 2001US-0300933P.  
XX  
(UYNC-) UNIV NORTH CAROLINA STATE.  
XX  
Martin LD, Adler KB, Li Y;  
XX  
WPI; 2003-278239/27.  
XX  
Method of regulating inflammation comprises administering a composition  
comprising a MANS peptide or an active fragment thereof.  
XX  
Disclosure; Page 26; 54pp; English.  
XX  
The present sequence represents a scrambled myristoylated N-terminal  
sequence (MANS) peptide. It is used as a control peptide, in the method  
of the invention. The MANS peptide is identical to the first 24 amino  
acids of MARCKS, and mediates insertion of MARCKS into membranes. The  
MANS peptide inhibits both mucus secretion and inflammatory mediators.  
XX  
The specification describes a method of regulating inflammation. The

CC method comprises administering a composition comprising a MAN5 peptide.  
 CC The method is useful for treating inflammation caused by respiratory  
 CC diseases (e.g. asthma, chronic bronchitis and chronic obstructive  
 CC pulmonary disease (COPD), bowel diseases (e.g. irritable bowel syndrome,  
 CC Crohn's disease and ulcerative colitis), skin diseases (e.g. rosacea,  
 CC eczema, psoriasis and severe acne), autoimmune diseases and pain  
 CC syndromes, arthritis and cystic fibrosis  
 CC  
 XX Sequence 24 AA;  
 SQ  
 Query Match 100.0%; Score 115; DB 6; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 6.8e-10;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GTAPAAEGAGAEVVRASAEAKQAF 24  
 |||||  
 DB 1 GTAPAAEGAGAEVVRASAEAKQAF 24  
 |||||  
 RESULT 2  
 ADD23833  
 ID ADD23833 standard; peptide; 18 AA.  
 AC  
 XX ADD23833;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Breast cancer membrane protein (BCMP) peptide SEQ ID NO:660.  
 XX  
 KW breast cancer; screening; diagnosis; breast cancer therapy;  
 KW breast cancer membrane protein; BCMP; cytostatic; vaccine; human.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO2003087831-A2.  
 XX  
 PD 23-OCT-2003.  
 XX  
 PF 10-APR-2003; 2003WO-GB001559.  
 XX  
 PR 11-APR-2002; 2002GB-00008331.  
 XX  
 PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.  
 XX  
 PI Hudson LJ, Stamps AC, Terrett JA;  
 XX  
 WPI; 2003-845381/78.  
 DR  
 PT Screening, diagnosing and/or treating breast cancer by detecting a change  
 PT in expression or activity of a breast cancer membrane protein (BCMP)  
 PT polypeptide or encoding nucleic acid molecule.  
 XX  
 PS Example; SEQ ID NO 660; 81pp; English.  
 XX  
 CC The present invention describes a method of screening for and/or  
 CC diagnosing breast cancer in a subject, and/or monitoring the  
 CC effectiveness of breast cancer therapy. The method comprises detecting  
 CC and/or quantifying in a biological sample obtained from the subject a  
 CC breast cancer membrane protein (BCMP) polypeptide and a nucleic acid  
 CC molecule. Also described: (1) an antibody, its functionally-active  
 CC fragment, derivative or analogue, that specifically binds to one or more  
 CC of the BCMP polypeptide; (2) a diagnostic kit comprising a capture  
 CC reagent specific for an BCMP polypeptide, reagents and instructions for  
 CC use; (3) a method for screening for anti-breast cancer agents that  
 CC interact with the BCMP polypeptide, comprising contacting the polypeptide  
 CC with a candidate agent, and determining whether or not the candidate  
 CC agent interacts with the polypeptide; (4) a method for screening for anti  
 CC -breast cancer agents that modulate the expression or activity of an BCMP  
 CC polypeptide or the nucleic acid molecule cited above, comprising  
 CC comparing the expression or activity of the polypeptide or nucleic acid  
 CC molecule, in the presence and absence of a candidate agent or in the  
 CC presence of a control agent, and determining whether the candidate agent  
 CC causes the expression or activity of the polypeptide or nucleic acid

CC molecule to change; and (5) an agent identified by the method of (3) or  
 CC (4), which interacts with the polypeptide or causes the expression or  
 CC activity of the polypeptide, or the expression of the nucleic acid  
 CC molecule to change. BCMPs have cytostatic activities, and can be used in  
 CC vaccines. The BCMP polypeptide, nucleic acid molecule, antibody, agent or  
 CC their derivatives, are useful in the manufacture of a medicament for the  
 CC treatment of breast cancer, where the composition is a vaccine. The  
 CC present sequence represents a BCMP peptide which is used in the  
 CC exemplification of the present invention.  
 CC  
 XX Sequence 18 AA;  
 SQ  
 Query Match 38.3%; Score 44; DB 7; Length 18;  
 Best Local Similarity 53.3%; Pred. No. 14;  
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 GTAPAAEGAGAEVVR 15  
 |||||  
 DB 4 GSAPAGEGEGVTQOR 18  
 |||||

RESULT 3  
 ABP62492  
 ID ABP62492 standard; peptide; 24 AA.  
 AC  
 XX ABP62492;  
 XX  
 DT 10-OCT-2002 (first entry)  
 XX  
 DE Human immunopeptide to HCV E2 glycoprotein framework sequence #31.  
 XX  
 KW Virucide; human; immunopolypeptide; immunopeptide; envelope glycoprotein;  
 KW nonstructural protein; hepatitis C virus; HCV; E2 glycoprotein;  
 KW NS3 protein; viral infection.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200259340-A1.  
 XX  
 PD 01-AUG-2002.  
 XX  
 PF 25-JAN-2002; 2002WO-US002303.  
 XX  
 PR 26-JAN-2001; 2001US-0264451P.  
 XX  
 PA (SCRI ) SCRIPPS RES INST.  
 XX  
 PI Maruyama T, Jones IM, Burton DR, Fox RI;  
 XX  
 WPI; 2002-599801/64.  
 DR  
 PT New human immunopolypeptide with binding specificity for certain envelope  
 PT glycoproteins and nonstructural proteins of hepatitis C virus (HCV), for  
 PT diagnosing or treating patients having or suspected of having HCV  
 PT infection.  
 XX  
 PS Claim 4; Fig 17; 308pp; English.  
 XX  
 CC The present invention relates to human immunopolypeptides, produced by a  
 CC phage transfected cell library. The present sequence is one such  
 CC immunopolypeptide. The immunopolypeptides have binding specificity for  
 CC envelope glycoprotein E2 and nonstructural protein NS3 of hepatitis C  
 CC virus (HCV). E2 glycoprotein is believed to be responsible for target  
 CC cell binding and contains neutralising epitopes, while NS3 is thought to  
 CC be involved in the replication of HCV. The immunopolypeptides are useful  
 CC for diagnosing and treating a patient having or suspected to be having  
 CC HCV infection  
 CC  
 XX Sequence 24 AA;  
 SQ  
 Query Match 33.0%; Score 38; DB 5; Length 24;  
 Best Local Similarity 53.3%; Pred. No. 1.5e+02;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 7 EGAGAEVKKRASAEAK 21  
 Db :|||||:  
 2 EQSGAEVKKPGASVK 16

RESULT 4  
 ABP12991  
 ID ABP12991 standard; peptide; 8 AA.

XX AC ABP12991;  
 XX DT 11-SEP-2003 (revised)  
 XX DT 15-JUL-2002 (first entry)  
 XX DE HIV A02 super motif nef peptide #4.

XX KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;  
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;  
 KW vaccine; HIV infection; immunisation; virucide.

XX OS Human immunodeficiency virus 1.

XX PN WO200124810-A1.

XX PD 12-APR-2001.

XX PF 05-OCT-2000; 2000WO-US027766.

XX PR 05-OCT-1999; 99US-00412863.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
 PI Baker DM, Celis E, Kubo RT, Grey HM;

XX WPI; 2001-354887/37.

XX Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
 PT peptide groups, useful for vaccinating against HIV-1.

XX Claim 32; Page 134; 448pp; English.

XX The present invention describes a composition (I) comprising a prepared  
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may  
 CC be used for immunising subjects against HIV-1 infections. The use of  
 CC group-based vaccines has several advantages over traditional vaccines,  
 CC particularly when compared to the use of whole antigens in vaccine  
 CC compositions. There is evidence that the immune response to whole  
 CC antigens is directed largely toward variable regions of the antigen,  
 CC allowing for immune escape due to mutations. The groups for inclusion in  
 CC an group-based vaccine may be selected from conserved regions of viral or  
 CC tumour-associated antigens, which therefore reduces the likelihood of  
 CC escape mutants. Furthermore, immunosuppressive groups that may be present  
 CC in whole antigens can be avoided with the use of group-based vaccines. An  
 CC additional advantage of an group-based vaccine approach is the ability to  
 CC combine selected groups (CTL and HTL), and further, to modify the  
 CC composition of the groups, achieving, for example, enhanced  
 CC immunogenicity. Accordingly, the immune response can be modulated, as  
 CC appropriate, for the target disease. Similar engineering of the response  
 CC is not possible with traditional approaches. ABP11501 to ABP25412  
 CC represent peptide sequences used in the exemplification of the present  
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)

XX Sequence 8 AA;

Query Match 31.3%; Score 36; DB 4; Length 8;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAAGAGAGA 11

Db :|||||:  
 1 PAAEGVGA 8

RESULT 5  
 ABP20973  
 ID ABP20973 standard; peptide; 8 AA.

XX AC ABP20973;

XX DT 11-SEP-2003 (revised)  
 XX DT 15-JUL-2002 (first entry)

XX DE HIV A03 motif nef peptide #123.

XX KW HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;  
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;  
 KW vaccine; HIV infection; immunisation; virucide.

XX OS Human immunodeficiency virus 1.

XX PN WO200124810-A1.

XX PD 12-APR-2001.

XX PF 05-OCT-2000; 2000WO-US027766.

XX PR 05-OCT-1999; 99US-00412863.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
 PI Baker DM, Celis E, Kubo RT, Grey HM;

XX WPI; 2001-354887/37.

XX Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
 PT peptide groups, useful for vaccinating against HIV-1.

XX Claim 32; Page 299; 448pp; English.

XX The present invention describes a composition (I) comprising a prepared  
 CC human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may  
 CC be used for immunising subjects against HIV-1 infections. The use of  
 CC group-based vaccines has several advantages over traditional vaccines,  
 CC particularly when compared to the use of whole antigens in vaccine  
 CC compositions. There is evidence that the immune response to whole  
 CC antigens is directed largely toward variable regions of the antigen,  
 CC allowing for immune escape due to mutations. The groups for inclusion in  
 CC an group-based vaccine may be selected from conserved regions of viral or  
 CC tumour-associated antigens, which therefore reduces the likelihood of  
 CC escape mutants. Furthermore, immunosuppressive groups that may be present  
 CC in whole antigens can be avoided with the use of group-based vaccines. An  
 CC additional advantage of an group-based vaccine approach is the ability to  
 CC combine selected groups (CTL and HTL), and further, to modify the  
 CC composition of the groups, achieving, for example, enhanced  
 CC immunogenicity. Accordingly, the immune response can be modulated, as  
 CC appropriate, for the target disease. Similar engineering of the response  
 CC is not possible with traditional approaches. ABP11501 to ABP25412  
 CC represent peptide sequences used in the exemplification of the present  
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)

XX Sequence 8 AA;

Query Match 31.3%; Score 36; DB 4; Length 8;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAAEGVGA 11

Db :|||||:  
 1 PAAEGVGA 8



DT 15-JUL-2002 (first entry)  
 XX HIV B58 super motif nef peptide #19.  
 DE HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;  
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;  
 KW vaccine; HIV infection; immunisation; virucide.  
 XX Human immunodeficiency virus 1.  
 OS WO200124810-A1.  
 XX PN 12-APR-2001.  
 XX PD 05-OCT-2000; 2000WO-US027766.  
 XX PF 05-OCT-1999; 99US-00412863.  
 XX PR (EPIM-) EPIMUNE INC.  
 XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
 PI Baker DM, Celis E, Kubo RT, Grey HM;  
 XX WPI; 2001-354887/37.  
 DR Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
 XX peptide groups, useful for vaccinating against HIV-1.  
 PT Claim 32; Page 237; 448pp; English.  
 PS The present invention describes a composition (I) comprising a prepared  
 XX human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may  
 CC be used for immunising subjects against HIV-1 infections. The use of  
 CC group-based vaccines has several advantages over traditional vaccines,  
 CC particularly when compared to the use of whole antigens in vaccine  
 CC compositions. There is evidence that the immune response to whole  
 CC antigens is directed largely toward variable regions of the antigen,  
 CC allowing for immune escape due to mutations. The groups for inclusion in  
 CC an group-based vaccine may be selected from conserved regions of viral or  
 CC tumour-associated antigens, which therefore reduces the likelihood of  
 CC escape mutants. Furthermore, immunosuppressive groups that may be present  
 CC in whole antigens can be avoided with the use of group-based vaccines. An  
 CC additional advantage of an group-based vaccine approach is the ability to  
 CC combine selected groups (CTL and HTL), and further, to modify the  
 CC composition of the groups, achieving, for example, enhanced  
 CC immunogenicity. Accordingly, the immune response can be modulated, as  
 CC appropriate, for the target disease. Similar engineering of the response  
 CC is not possible with traditional approaches. ABP11501 to ABP25412  
 CC represent peptide sequences used in the exemplification of the present  
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)  
 XX Sequence 9 AA;  
 SQ Query Match 31.3%; Score 36; DB 4; Length 9;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 PAACGAGA 11  
 Db 1 PAAEGVGA 8  
 RESULT 9  
 ABP13017  
 ID ABP13017 standard; peptide; 9 AA.  
 XX AC ABP13017;  
 XX 11-SEP-2003 (revised)  
 DT 15-JUL-2002 (first entry)  
 XX

DE HIV A02 super motif nef peptide #30.  
 XX HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;  
 KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;  
 KW vaccine; HIV infection; immunisation; virucide.  
 XX Human immunodeficiency virus 1.  
 OS WO200124810-A1.  
 XX PN 12-APR-2001.  
 XX PD 05-OCT-2000; 2000WO-US027766.  
 XX PF 05-OCT-1999; 99US-00412863.  
 XX PR (EPIM-) EPIMUNE INC.  
 XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
 PI Baker DM, Celis E, Kubo RT, Grey HM;  
 XX WPI; 2001-354887/37.  
 DR Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
 XX peptide groups, useful for vaccinating against HIV-1.  
 PT Claim 32; Page 135; 448pp; English.  
 PS The present invention describes a composition (I) comprising a prepared  
 XX human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
 CC sequence selected from 51 defined amino acid sequences (ABL25347 to  
 CC ABP25397). (I) has virucide activity and can be used in vaccines. (I) may  
 CC be used for immunising subjects against HIV-1 infections. The use of  
 CC group-based vaccines has several advantages over traditional vaccines,  
 CC particularly when compared to the use of whole antigens in vaccine  
 CC compositions. There is evidence that the immune response to whole  
 CC antigens is directed largely toward variable regions of the antigen,  
 CC allowing for immune escape due to mutations. The groups for inclusion in  
 CC an group-based vaccine may be selected from conserved regions of viral or  
 CC tumour-associated antigens, which therefore reduces the likelihood of  
 CC escape mutants. Furthermore, immunosuppressive groups that may be present  
 CC in whole antigens can be avoided with the use of group-based vaccines. An  
 CC additional advantage of an group-based vaccine approach is the ability to  
 CC combine selected groups (CTL and HTL), and further, to modify the  
 CC composition of the groups, achieving, for example, enhanced  
 CC immunogenicity. Accordingly, the immune response can be modulated, as  
 CC appropriate, for the target disease. Similar engineering of the response  
 CC is not possible with traditional approaches. ABP11501 to ABP25412  
 CC represent peptide sequences used in the exemplification of the present  
 CC invention. (Updated on 11-SEP-2003 to standardise OS field)  
 XX Sequence 9 AA;  
 SQ Query Match 31.3%; Score 36; DB 4; Length 9;  
 Best Local Similarity 87.5%; Pred. No. 1.4e+06;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 PAACGAGA 11  
 Db 1 PAAEGVGA 8  
 RESULT 10  
 ADD57289  
 ID ADD57289 standard; peptide; 10 AA.  
 XX AC ADD57289;  
 XX 15-JAN-2004 (first entry)  
 DT HLA binding epitope #509.  
 DE Virucide; cytostatic; gene therapy; vaccine; epitope; cytotoxic T cell;  
 KW

KW MHC class I; CTL; HTL; A2-restricted cytotoxic lymphocyte; HLA;  
KW viral disease; cancer.  
XX Unidentified.  
XX WO2003040165-A2.  
XX 15-MAY-2003.  
XX 18-OCT-2001; 2001WO-US051650.  
XX 19-OCT-2000; 2000US-0242350P.  
XX 20-APR-2001; 2001US-0285624P.  
XX (EPIM-) EPIMUNE INC.  
XX Sette A, Sidney J, Southwood S;  
XX WPI; 2003-441519/41.  
XX  
XX New composition comprising at least one peptide having allele-specific  
XX binding motifs for HLA, useful for preventing, treating or diagnosing  
XX viral diseases and cancer.  
XX  
XX Claim 1; Page 52-379; 382pp; English.  
XX  
XX The invention relates to a composition comprising at least one peptide  
XX having an isolated, prepared epitope selected from any of the sequences  
XX from 30 lists given in the specification. Also disclosed is a method for  
XX inducing a cytotoxic T cell response against a pre-selected antigen in a  
XX patient expressing a specific MHC class I allele by contacting cytotoxic  
XX T cells from the patient with the composition cited above. The  
XX composition comprises an epitope that is joined by an amino acid linker.  
XX The epitope is admixed or joined to a CTL or HTL epitope. The epitope is  
XX bound to an HLA molecule on the antigen-presenting cell, where when an A2  
XX -restricted cytotoxic lymphocyte (CTL) is present, a receptor of the CTL  
XX binds to a complex of the HLA molecule and the epitope. Specifically  
XX claimed are peptides having allele-specific binding motifs for HLA. The  
XX compositions and methods are useful for preventing, treating or  
XX diagnosing viral diseases and cancer. The peptide epitopes are useful as  
XX diagnostic agents for evaluating immune responses, for making antibodies  
XX and for evaluating efficacy of a vaccine. Sequences given in ADD56781-  
XX ADD65275 represent epitopes of the invention as given in Tables 2-31.  
XX  
XX Sequence 10 AA;  
SQ  
Query Match 31.3%; Score 36; DB 7; Length 10;  
Best Local Similarity 87.5%; Pred. No. 1.1e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 4 PAAEGGGA 11  
Db 3 PAAEGGGA 10  
RESULT 11  
ABP13067  
ID ABP13067 standard; peptide; 11 AA.  
XX  
XX ABP13067;  
XX  
XX 11-SEP-2003 (revised)  
XX 15-JUL-2002 (first entry)  
XX HIV A02 super motif nef peptide #80.  
XX HIV; HIV-1; human immunodeficiency virus; env; pol; gag; nef; vpr; vpu;  
KW vif; tat; cytotoxic T lymphocyte; CTL; immune response; epitope; antigen;  
KW vaccine; HIV infection; immunisation; virucide.  
XX  
XX Human immunodeficiency virus 1.  
XX  
XX WO200124810-A1.  
PN

XX 12-APR-2001.  
XX  
XX 05-OCT-2000; 2000WO-US027766.  
XX  
XX 05-OCT-1999; 99US-00412863.  
XX (EPIM-) EPIMUNE INC.  
XX  
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
XX Baker DM, Celis E, Kubo RT, Grey HM;  
XX WPI; 2001-354887/37.  
XX Vaccine compositions comprising human immunodeficiency virus-1 (HIV-1)  
XX peptide groups, useful for vaccinating against HIV-1.  
XX  
XX Claim 32; Page 136; 448pp; English.  
XX  
XX The present invention describes a composition (I) comprising a prepared  
XX human immunodeficiency virus-1 (HIV-1) group comprising an amino acid  
XX sequence selected from 51 defined amino acid sequences (ABL25347 to  
XX ABP25397). (I) has virucide activity and can be used in vaccines. (I) may  
XX be used for immunising subjects against HIV-1 infections. The use of  
XX group-based vaccines has several advantages over traditional vaccines,  
XX particularly when compared to the use of whole antigens in vaccine  
XX compositions. There is evidence that the immune response to whole  
XX antigens is directed largely toward variable regions of the antigen,  
XX allowing for immune escape due to mutations. The groups for inclusion in  
XX an group-based vaccine may be selected from conserved regions of viral or  
XX tumour-associated antigens, which therefore reduces the likelihood of  
XX escape mutants. Furthermore, immunosuppressive groups that may be present  
XX in whole antigens can be avoided with the use of group-based vaccines. An  
XX additional advantage of an group-based vaccine approach is the ability to  
XX combine selected groups (CTL and HTL), and further, to modify the  
XX composition of the groups, achieving, for example, enhanced  
XX immunogenicity. Accordingly, the immune response can be modulated, as  
XX appropriate, for the target disease. Similar engineering of the response  
XX is not possible with traditional approaches. ABP11501 to ABP25412  
XX represent peptide sequences used in the exemplification of the present  
XX invention. (Updated on 11-SEP-2003 to standardise OS field)  
XX  
XX Sequence 11 AA;  
SQ  
Query Match 31.3%; Score 36; DB 4; Length 11;  
Best Local Similarity 87.5%; Pred. No. 1.3e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 4 PAAEGGGA 11  
Db 4 PAAEGGGA 11  
RESULT 12  
ADD57290  
ID ADD57290 standard; peptide; 11 AA.  
XX  
XX ADD57290;  
XX  
XX 15-JAN-2004 (first entry)  
XX  
XX HLA binding epitope #510.  
XX Virucide; cytostatic; gene therapy; vaccine; epitope; cytotoxic T cell;  
KW MHC class I; CTL; HTL; A2-restricted cytotoxic lymphocyte; HLA;  
KW viral disease; cancer.  
XX Unidentified.  
XX  
XX WO2003040165-A2.  
XX  
XX 15-MAY-2003.  
XX



PF 18-OCT-2001; 2001WO-US051650.  
 XX  
 PR 19-OCT-2000; 2000US-0242350P.  
 PR 20-APR-2001; 2001US-0285624P.  
 XX  
 PA (EPIM-) EPIMMUNE INC.  
 XX  
 PI Sette A, Sidney J, Southwood S;  
 XX  
 DR WPI; 2003-441519/41.  
 XX  
 XX  
 PT New composition comprising at least one peptide having allele-specific  
 PT binding motifs for HLA, useful for preventing, treating or diagnosing  
 PT viral diseases and cancer.  
 XX  
 XX Claim 1; Page 52-379; 382pp; English.  
 PS  
 PS The invention relates to a composition comprising at least one peptide  
 CC having an isolated, prepared epitope selected from any of the sequences  
 CC from 30 lists given in the specification. Also disclosed is a method for  
 CC inducing a cytotoxic T cell response against a pre-selected antigen in a  
 CC patient expressing a specific MHC class I allele by contacting cytotoxic  
 CC T cells from the patient with the composition cited above. The  
 CC composition comprises an epitope that is joined by an amino acid linker.  
 CC The epitope is admixed or joined to a CTL or HTL epitope. The epitope is  
 CC bound to an HLA molecule on the antigen-presenting cell, where when an A2  
 CC -restricted cytotoxic lymphocyte (CTL) is present, a receptor of the CTL  
 CC binds to a complex of the HLA molecule and the epitope. Specifically  
 CC claimed are peptides having allele-specific binding motifs for HLA. The  
 CC compositions and methods are useful for preventing, treating or  
 CC diagnosing viral diseases and cancer. The peptide epitopes are useful as  
 CC diagnostic agents for evaluating immune responses, for making antibodies  
 CC and for evaluating efficacy of a vaccine. Sequences given in ADP56781-  
 CC ADD65275 represent epitopes of the invention as given in Tables 2-31.  
 XX  
 SQ Sequence 11 AA;  
 Query Match 31.3%; Score 36; DB 7; Length 11;  
 Best Local Similarity 87.5%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 PAAEGAGA 11  
 DB 3 PAAEGVGA 10  
 RESULT 13  
 AAR43045  
 ID AAR43045 standard; protein; 20 AA.  
 XX  
 AC AAR43045;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 10-MAR-2003 (revised)  
 DT 24-MAY-1994 (first entry)  
 XX  
 XX N-terminal fragment.  
 DE  
 XX  
 KW Asparaginylendopeptidase; Canavalia ensiformis; seed; L-asparagine;  
 KW primer; PCR; protein fragmentation; peptide synthesis.  
 XX  
 OS Canavalia ensiformis.  
 XX  
 FN JP05276960-A.  
 XX  
 PD 26-OCT-1993.  
 XX  
 PF 07-AUG-1992; 92JP-00231602.  
 XX  
 XX 07-FEB-1992; 92JP-00056023.  
 XX  
 PA (TAKI) TAKARA SHUZO CO LTD.  
 XX

DR WPI; 1993-373587/47.  
 XX  
 PT New gene for encoding asparaginyl endo-peptidase - comprises 8 specified  
 PT DNA sequences.  
 XX  
 PS Disclosure; Page 30; 35pp; Japanese.  
 XX  
 CC A gene encoding asparaginylendopeptidase is claimed. 8 DNA sequences are  
 CC given (AAQ50559-66). The enzyme is a protease derived from a seed of  
 CC Canavalia ensiformis which selectively hydrolyses C-terminus amide bond  
 CC of L-asparagine residue (see AAR43033 and AAR43041). The enzyme is useful  
 CC for protein fragmentation and enzymatic peptide synthesis. The primers  
 CC given in AAQ50567-68, AAQ50576-77 and AAQ50583-90 were used in the  
 CC isolation of the fragments given in AAQ50569-75 and AAQ50578-79, by PCR.  
 CC (Updated on 10-MAR-2003 to add missing OS field.) (Updated on 25-MAR-2003  
 CC to correct PA field.)  
 XX  
 SQ Sequence 20 AA;  
 Query Match 31.3%; Score 36; DB 2; Length 20;  
 Best Local Similarity 47.4%; Pred. No. 2.4e+02;  
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
 QY 4 PAAEGAGAEVKRAAEAKQ 22  
 DB 2 PAAAKRKEAEELAAATAEQ 20  
 RESULT 14  
 AAR95258  
 ID AAR95258 standard; protein; 22 AA.  
 XX  
 AC AAR95258;  
 XX  
 DT 09-NOV-1996 (first entry)  
 XX  
 DE Vector encoded amino acids used to increase protease yield.  
 XX  
 KW Protease; restriction protease; cleavage; proteolytic domain; protein.  
 XX  
 OS Synthetic.  
 XX  
 FN WO9616167-A1.  
 XX  
 PD 30-MAY-1996.  
 XX  
 PF 16-NOV-1995; 95WO-NO000213.  
 XX  
 PR 17-NOV-1994; 94NO-00004411.  
 XX  
 PA (HAVA/) HAVARSTEIN L S.  
 PA (NESI/) NES I F.  
 XX  
 PI Havarstein LS, Nes IF;  
 XX  
 DR WPI; 1996-268604/27.  
 XX  
 PT Protease(s) with highly specific cleavage activity - have proteolytic  
 PT domains derived from N-terminus of ABC transporters, pref. 150 N-terminal  
 PT amino acids of LagD.  
 XX  
 PS Claim 10; Page 22; 43pp; English.  
 XX  
 CC Proteases derived from ABC-transporters containing N-terminal proteolytic  
 CC domains are restriction proteases that can be used for the cleavage of  
 CC proteins at specific amino acid cleavage sites. These 22 vector encoded  
 CC amino acids were added to the C-terminus of the subcloned protease  
 CC sequence to substantially increase the yield of the recombinant ABC  
 CC protease whilst not reducing its enzymatic activity of the protease  
 XX  
 SQ Sequence 22 AA;  
 Query Match 31.3%; Score 36; DB 2; Length 22;

Best Local Similarity 47.4%; Pred. No. 2.7e+02;  
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 PAAEAGAGAEVKKRASAEEAKQ 22  
||| | | : | : | : |  
Db 4 PAANKARKEAEELAAATAEQ 22

RESULT 15  
AAU99665  
ID AAU99665 standard; peptide; 22 AA.  
XX  
AC AAU99665;  
XX  
DT 07-OCT-2002 (first entry)  
XX  
DE PET15b encoded C-terminal extension peptide.  
XX  
KW PET15b; C terminal extension; adenovirus binding domain; D1; D2; sd1;  
KW coxsackievirus and adenovirus receptor; CAR; protein therapy; virucide;  
KW adenovirus infection; coxsackievirus infection; ocular infection;  
KW upper respiratory tract infection; gastrointestinal infection;  
KW chromosome 21; virus attachment protein; Adenovirus 12; knob;  
KW receptor-knob interface.  
XX  
OS Synthetic.  
XX  
PN US6395875-B1.  
XX  
PD 28-MAY-2002.  
XX  
PF 25-JAN-1999; 99US-00236423.  
XX  
PR 25-JAN-1999; 99US-00236423.  
XX  
PA (BROO-) BROOKHAVEN SCI ASSOC LLC.  
XX  
PI Freimuth PI;  
XX  
DR WPI; 2002-556105/59.  
XX  
PT New polypeptide for preparing a composition to treat ocular, upper  
PT respiratory or gastrointestinal infection, comprises an adenovirus  
PT binding domain D1 of human coxsackievirus and an adenovirus receptor  
PT protein.  
XX  
PS Claim 5; Fig 1; 13pp; English.

CC The invention relates to an isolated polypeptide consisting essentially  
CC of an amino acid sequence which corresponds to adenovirus binding domain  
CC D1 of human coxsackievirus and adenovirus receptor (CAR) protein. Also  
CC included are an isolated polypeptide consisting essentially of an amino  
CC acid sequence which corresponds to extracellular domains D1 and D2 of  
CC human CAR protein, where the amino acid sequence corresponds to residues  
CC 20-237 of human pre-CAR protein with the exception of two amino acid  
CC substitutions, Leu20Met and Ser21Gly and an isolated fusion protein (sd1)  
CC comprising a first amino acid sequence which corresponds to amino acids  
CC 20-144 of human pre-CAR protein (D1) fused in frame to a second amino  
CC acid sequence comprising the peptide appearing as AAU99665 where the  
CC second amino acid sequence facilitates folding of D1 into a functional,  
CC soluble domain when expressed in bacteria and is located C-terminal to  
CC the first amino acid sequence, being encoded by the vector PET15b. CAR D1  
CC functions as an antiviral agent by inhibiting viral infection of a cell.  
CC CAR D1 is useful for treating a patient infected with virus e.g.,  
CC adenovirus of group A, adenovirus of group C and coxsackievirus subgroup  
CC B. D1 is also useful for inhibiting the progression or spread of  
CC infection and for the preparation of a therapeutic composition for  
CC treating an ocular, upper respiratory or gastrointestinal infection,  
CC respectively. D1 is also useful for identifying and characterising  
CC molecules which bind CAR through the D1 domain to study the infection  
CC process and to develop new therapeutics. D1 is also utilised in binding  
CC assays for identification and further characterisation of D1. D1 is also  
CC utilised in binding assays to screen for antiviral compounds which

CC interfere specifically with the interaction between the D1 polypeptide of  
CC the human CAR protein and a viral attachment protein which binds to the  
CC D1 polypeptide. D1 is useful for structural determination of virus  
CC attachment proteins (e.g., adenovirus 12 (Ad12) knob, complexed to D1 or  
CC D1/D2) which facilitates the identification of residues that form the  
CC receptor-knob interface. The gene for human CAR is located on chromosome  
CC 21. The present sequence is the PET15b encoded C-terminal extension  
XX peptide  
XX  
SQ Sequence 22 AA;  
Query Match 31.3%; Score 36; DB 5; Length 22;  
Best Local Similarity 47.4%; Pred. No. 2.7e+02;  
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 PAAEAGAGAEVKKRASAEEAKQ 22  
||| | | : | : | : |  
Db 4 PAANKARKEAEELAAATAEQ 22

Search completed: October 7, 2004, 17:53:34  
Job time : 88.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:54:08 ; Search time 83.5 Seconds  
(without alignments)  
92.493 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115  
Sequence: 1 GTAPAEAGAGAEVKRASAERAKQAF 24

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 284360

Minimum DB seq length: 0  
Maximum DB seq length: 24

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/ECT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/ECTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09D\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10E\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	100.0	24	12 US-10-180-753-2	Sequence 2, Appli
2	115	100.0	24	15 US-10-802-644-2	Sequence 2, Appli
3	38	33.0	21	15 US-10-313-986-547	Sequence 547, App
4	36	31.3	22	13 US-10-155-282-1	Sequence 1, Appli
5	36	31.3	22	14 US-10-199-290-1	Sequence 1, Appli
6	36	31.3	22	14 US-10-218-419-1	Sequence 1, Appli
7	36	31.3	22	14 US-10-037-243-20	Sequence 20, Appl
8	36	31.3	22	14 US-10-037-243-42	Sequence 42, Appl
9	36	31.3	23	14 US-10-037-243-4	Sequence 4, Appli
10	34	29.6	14	15 US-10-394-980-142	Sequence 142, App
11	34	29.6	20	14 US-10-029-386-27605	Sequence 27605, A
12	33.5	29.1	10	10 US-09-572-404B-1296	Sequence 1296, Ap
13	33.5	29.1	10	10 US-09-572-404B-3822	Sequence 3822, Ap
14	33	28.7	19	8 US-08-344-824-276	Sequence 276, App
15	33	28.7	14	10 US-09-824-438-16	Sequence 16, Appl

16	33	28.7	16	16 US-10-433-452A-54	Sequence 54, Appl
17	33	28.7	21	12 US-10-169-613-13	Sequence 13, Appl
18	33	28.7	21	12 US-10-169-613-25	Sequence 25, Appl
19	32	27.8	15	15 US-10-149-138-3748	Sequence 3748, Ap
20	32	27.8	15	16 US-10-149-138-3748	Sequence 3748, Ap
21	32	27.8	18	10 US-09-931-325A-17	Sequence 17, Appl
22	32	27.8	18	10 US-09-931-325A-160	Sequence 160, App
23	32	27.8	18	10 US-09-930-915A-36	Sequence 36, Appl
24	32	27.8	18	10 US-09-930-915A-200	Sequence 200, App
25	32	27.8	18	14 US-10-082-014-59	Sequence 59, Appl
26	32	27.8	18	14 US-10-372-076-60	Sequence 60, Appl
27	32	27.8	18	16 US-10-806-006-36	Sequence 36, Appl
28	32	27.8	18	16 US-10-806-006-200	Sequence 200, App
29	32	27.8	18	16 US-10-677-074-60	Sequence 60, Appl
30	32	27.8	18	16 US-10-805-913-36	Sequence 36, Appl
31	32	27.8	18	16 US-10-805-913-200	Sequence 200, App
32	32	27.8	20	14 US-10-355-975-27	Sequence 27, Appl
33	32	27.8	20	16 US-10-803-622-185	Sequence 185, App
34	32	27.8	20	16 US-10-803-653-185	Sequence 185, App
35	32	27.8	21	10 US-09-931-325A-88	Sequence 88, Appl
36	32	27.8	21	10 US-09-930-915A-129	Sequence 129, App
37	32	27.8	21	14 US-10-082-014-244	Sequence 244, App
38	32	27.8	21	14 US-10-372-076-274	Sequence 274, App
39	32	27.8	21	16 US-10-806-006-129	Sequence 129, App
40	32	27.8	21	16 US-10-677-074-274	Sequence 274, App
41	32	27.8	21	16 US-10-805-913-129	Sequence 129, App
42	31	27.0	11	15 US-10-149-138-1495	Sequence 1495, Ap
43	31	27.0	11	16 US-10-149-138-1495	Sequence 1495, Ap
44	31	27.0	16	12 US-09-839-577A-2	Sequence 2, Appli
45	31	27.0	16	12 US-09-905-691-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1  
US-10-180-753-2  
; Sequence 2, Application US/10180753  
; Publication No. US20030013652A1  
; GENERAL INFORMATION:  
; APPLICANT: Martin, Linda  
; APPLICANT: Adler, Kenneth  
; APPLICANT: Li, Yuehua  
; TITLE OF INVENTION: BLOCKING PEPTIDE FOR INFLAMMATORY CELL SECRETION  
; FILE REFERENCE: 5051.574  
; CURRENT APPLICATION NUMBER: US/10/180,753  
; CURRENT FILING DATE: 2002-06-26  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 24  
; TYPE: PPT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic peptide  
; FEATURE:  
; NAME/KEY: MOD RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: N-terminal myristoylation  
US-10-180-753-2

Query Match 100.0%; Score 115; DB 12; Length 24;  
Best Local Similarity 100.0%; Pred. No. 4.3e-09;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTAPAEAGAGAEVKRASAERAKQAF 24  
|||||  
Db 1 GTAPAEAGAGAEVKRASAERAKQAF 24  
|||||

RESULT 2  
US-10-802-644-2  
; Sequence 2, Application US/10802644

```

; Publication No. US20040180836A1
; GENERAL INFORMATION:
; APPLICANT: Martin, Linda
; APPLICANT: Adler, Kenneth
; APPLICANT: Li, Yuehua
; TITLE OF INVENTION: BLOCKING PEPTIDE FOR INFLAMMATORY CELL SECRETION
; FILE REFERENCE: 5051.574
; CURRENT APPLICATION NUMBER: US/10/802,644
; CURRENT FILING DATE: 2004-03-17
; PRIOR APPLICATION NUMBER: US/10/180,753
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: N-terminal myristoylation
US-10-802-644-2

Query Match      100.0%; Score 115; DB 16; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.3e+09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTAPAAEGAGAEVVKRASAEAKQAF 24
Db 1 GTAPAAEGAGAEVVKRASAEAKQAF 24

RESULT 3
US-10-313-986-547
; Sequence 547, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 547
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-547

Query Match      33.0%; Score 38; DB 15; Length 21;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 9 AGAEVVKRASAEAKQA 23
Db 7 AGASIKIAPAEAPDA 21

RESULT 4
US-10-155-282-1
; Sequence 1, Application US/10155282
; Publication No. US20020151686A1
; GENERAL INFORMATION:
; APPLICANT: Freimuth, Paul
; TITLE OF INVENTION: RECOMBINANT SOLUBLE ADENOVIRUS RECEPTOR
; FILE REFERENCE: ENL-2007
; CURRENT APPLICATION NUMBER: US/10/155,282
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US/09/236,423
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: CLONING VECTOR
; OTHER INFORMATION: ENCODED SEQUENCES
US-10-155-282-1

Query Match      31.3%; Score 36; DB 13; Length 22;
Best Local Similarity 47.4%; Pred. No. 3.7e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 4 PAAEGAGAEVVKRASAEAKQ 22
Db 4 PAANKARKEAEELAAATAEQ 22

RESULT 5
US-10-199-290-1
; Sequence 1, Application US/10199290
; Publication No. US20030027338A1
; GENERAL INFORMATION:
; APPLICANT: Brookhaven Science Associates
; APPLICANT: Freimuth, Paul I
; TITLE OF INVENTION: Structure of Adenovirus Bound to Cellular Receptor CAR
; FILE REFERENCE: BSA 02-24
; CURRENT APPLICATION NUMBER: US/10/199,290
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: 09/389,603
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 09/236,423
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cloning vector encoded sequence
; OTHER INFORMATION: nces
US-10-199-290-1

Query Match      31.3%; Score 36; DB 14; Length 22;
Best Local Similarity 47.4%; Pred. No. 3.7e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 4 PAAEGAGAEVVKRASAEAKQ 22
Db 4 PAANKARKEAEELAAATAEQ 22

RESULT 6
US-10-218-419-1
; Sequence 1, Application US/10218419
; Publication No. US20030077813A1
; GENERAL INFORMATION:
; APPLICANT: Brookhaven Science Associates
; APPLICANT: Freimuth, Paul I
; TITLE OF INVENTION: Structure of Adenovirus Bound to Cellular Receptor CAR
; FILE REFERENCE: BSA 02-24
; CURRENT APPLICATION NUMBER: US/10/218,419
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 09/389,603
; PRIOR FILING DATE: 1999-09-03
; 
```

Cloning vector encoded sequence

```
; PRIOR APPLICATION NUMBER: 09/236,423
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Cloning vector encoded sequence
; OTHER INFORMATION: nces
US-10-218-419-1

Query Match      31.3%; Score 36; DB 14; Length 22;
Best Local Similarity 47.4%; Pred. No. 3.7e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      4 PAAEGAGAEVVKRSASAEAKQ 22
      ||| | | | | | | | | | | | | | | |
Db      4 PAANKARKEAEALAAATAEQ 22

RESULT 7
US-10-037-243-20
; Sequence 20, Application US/10037243
; Publication No. US20030134352A1
; GENERAL INFORMATION:
; APPLICANT: Brookhaven Science Associates, LLC.
; APPLICANT: Freimuth, Paul I
; APPLICANT: Zhang, Yian-Biao
; APPLICANT: Howitt, Jason A
; TITLE OF INVENTION: Facilitating Protein Folding and Solubility by Use of Peptide Ext
; FILE REFERENCE: BSA 01-22
; CURRENT APPLICATION NUMBER: US/10/037,243
; CURRENT FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic T7A peptide
US-10-037-243-20

Query Match      31.3%; Score 36; DB 14; Length 22;
Best Local Similarity 47.4%; Pred. No. 3.7e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      4 PAAEGAGAEVVKRSASAEAKQ 22
      ||| | | | | | | | | | | | | | | |
Db      4 PAANKARKEAEALAAATAEQ 22

RESULT 8
US-10-037-243-42
; Sequence 42, Application US/10037243
; Publication No. US20030134352A1
; GENERAL INFORMATION:
; APPLICANT: Brookhaven Science Associates, LLC.
; APPLICANT: Freimuth, Paul I
; APPLICANT: Zhang, Yian-Biao
; APPLICANT: Howitt, Jason A
; TITLE OF INVENTION: Facilitating Protein Folding and Solubility by Use of Peptide Ext
; FILE REFERENCE: BSA 01-22
; CURRENT APPLICATION NUMBER: US/10/037,243
; CURRENT FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Encoded amino acids
US-10-037-243-42

Query Match      31.3%; Score 36; DB 14; Length 22;
Best Local Similarity 47.4%; Pred. No. 3.7e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      4 PAAEGAGAEVVKRSASAEAKQ 22
      ||| | | | | | | | | | | | | | | |
Db      4 PAANKARKEAEALAAATAEQ 22

RESULT 9
US-10-037-243-4
; Sequence 4, Application US/10037243
; Publication No. US20030134352A1
; GENERAL INFORMATION:
; APPLICANT: Brookhaven Science Associates, LLC.
; APPLICANT: Freimuth, Paul I
; APPLICANT: Zhang, Yian-Biao
; APPLICANT: Howitt, Jason A
; TITLE OF INVENTION: Facilitating Protein Folding and Solubility by Use of Peptide Ext
; FILE REFERENCE: BSA 01-22
; CURRENT APPLICATION NUMBER: US/10/037,243
; CURRENT FILING DATE: 2002-06-03
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic peptide
US-10-037-243-4

Query Match      31.3%; Score 36; DB 14; Length 23;
Best Local Similarity 47.4%; Pred. No. 3.9e+02;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      4 PAAEGAGAEVVKRSASAEAKQ 22
      ||| | | | | | | | | | | | | | | |
Db      5 PAANKARKEAEALAAATAEQ 23

RESULT 10
US-10-394-980-142
; Sequence 142, Application US/10394980
; Publication No. US20040005633A1
; GENERAL INFORMATION:
; APPLICANT: Vandekerckhove, Joel
; APPLICANT: Gevaert, Kris
; TITLE OF INVENTION: METHODS AND APPARATUS FOR GEL-FREE QUALITATIVE AND
; FILE REFERENCE: VBV-001
; CURRENT APPLICATION NUMBER: US/10/394,980
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: PCT/EP02/03368
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: US60/278,171
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: US60/318,749
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US60/323,999
; PRIOR FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 473
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 142
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
```

; OTHER INFORMATION: part of MK01\_HUMAN (P28482, mitogen-activated protein kinase 1)  
US-10-394-980-142

Query Match 29.6%; Score 34; DB 15; Length 14;  
Best Local Similarity 61.5%; Pred. No. 4.4e+02;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 APAAEGAGAEVKR 15  
Db 2 AAAAAGAGPEWVR 14

## RESULT 11

US-10-029-386-27605  
; Sequence 27605, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: AEOICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 27605  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL139224.1  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.76  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3  
US-10-029-386-27605

Query Match 29.6%; Score 34; DB 14; Length 20;  
Best Local Similarity 35.7%; Pred. No. 6.4e+02;  
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 TAPAAEGAGAEVKR 15  
Db 7 SAPSVSLGCRIRK 20

## RESULT 12

US-09-572-404B-1296  
; Sequence 1296, Application US/09572404B  
; Publication No. US20030078374A1  
; GENERAL INFORMATION:  
; APPLICANT: Proteom Ltd  
; TITLE OF INVENTION: Complementary peptide ligands from the human genome  
; FILE REFERENCE: Human patent  
; CURRENT APPLICATION NUMBER: US/09/572,404B  
; CURRENT FILING DATE: 2000-05-17  
; NUMBER OF SEQ ID NOS: 4203  
; SOFTWARE: ProtPatent version 1.0  
; SEQ ID NO 1296  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; OTHER INFORMATION: sequence located in FKHL8 OR FREAC4 at 91-100 and may interact w  
; OTHER INFORMATION: Sequence 1295 in this patent.  
US-09-572-404B-1296

Query Match 29.1%; Score 33.5; DB 10; Length 10;  
Best Local Similarity 81.8%; Pred. No. 3.6e+02;  
Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 GTAPAAEGAGA 11  
Db 1 GPAPAA-GAGA 10

## RESULT 13

US-09-572-404B-3822  
; Sequence 3822, Application US/09572404B  
; Publication No. US20030078374A1  
; GENERAL INFORMATION:  
; APPLICANT: Proteom Ltd  
; TITLE OF INVENTION: Complementary peptide ligands from the human genome  
; FILE REFERENCE: Human patent  
; CURRENT APPLICATION NUMBER: US/09/572,404B  
; CURRENT FILING DATE: 2000-05-17  
; NUMBER OF SEQ ID NOS: 4203  
; SOFTWARE: ProtPatent version 1.0  
; SEQ ID NO 3822  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; OTHER INFORMATION: sequence located in FKHL8 OR FREAC4 at 91-100 and may interact w  
; OTHER INFORMATION: Sequence 3823 in this patent.  
US-09-572-404B-3822

Query Match 29.1%; Score 33.5; DB 10; Length 10;  
Best Local Similarity 81.8%; Pred. No. 3.6e+02;  
Matches 9; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1 GTAPAAEGAGA 11  
Db 1 GPAPAA-GAGA 10

## RESULT 14

US-08-344-824-276  
; Sequence 276, Application US/08344824  
; Publication No. US20030152580A1  
; GENERAL INFORMATION:  
; APPLICANT: SETTE, Alessandro  
; APPLICANT: SIDNEY, John  
; TITLE OF INVENTION: HLA BINDING PEPTIDES AND THEIR USES  
; NUMBER OF SEQUENCES: 399  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: One Market Plaza, Steuart Street Tower, 20th  
; STREET: Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/344,824  
; FILING DATE: 23-NOV-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/278,634  
; FILING DATE: 21-JUL-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bastian, Kevin L.  
; REGISTRATION NUMBER: 34,774  
; REFERENCE/DOCKET NUMBER: 14137-80-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 276:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-08-344-824-276

Query Match 28.7%; Score 33; DB 8; Length 9;  
Best Local Similarity 75.0%; Pred. No. 1.2e+06;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAAREGAGA 11  
|||:|  
Db 2 PAADGVGA 9

RESULT 15  
US-09-824-438-16  
; Sequence 16, Application US/09824438  
; Publication No. US20030073621A1  
; GENERAL INFORMATION:  
; APPLICANT: Butler, Jon P.  
; APPLICANT: Hale, John E.  
; APPLICANT: Heath Jr., William F.  
; APPLICANT: Schoner, Brigitte E.  
; APPLICANT: Heiman, Mark L.  
; APPLICANT: Becker, Gerald W.  
; APPLICANT: Varshavsky, Alexander D.  
; TITLE OF INVENTION: Beta-lipotropin and Uses Thereof  
; FILE REFERENCE: X-12139  
; CURRENT APPLICATION NUMBER: US/09/824,438  
; CURRENT FILING DATE: 2001-10-29  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-824-438-16

Query Match 28.7%; Score 33; DB 10; Length 14;  
Best Local Similarity 66.7%; Pred. No. 6.1e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAAREGAGAE 12  
||:|  
Db 2 PADDGGAQ 10

Search completed: October 7, 2004, 18:10:48  
Job time : 85.5 secs

**This Page Blank (uspio)**



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2004, 17:50:43 ; Search time 22 Seconds  
(without alignments)  
56.319 Million cell updates/sec

Title: US-10-802-644-2

Perfect score: 115

Sequence: 1 GTAPAAEGAGAEVVKRAAEAKQAF 24

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 189146

Minimum DB seq length: 0

Maximum DB seq length: 24

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*

4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*

5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*

6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	33.0	20	US-08-053-451B-171	Sequence 171, Appl
2	36	31.3	21	US-09-319-570A-4	Sequence 4, Appli
3	36	31.3	22	US-09-236-423-1	Sequence 1, Appli
4	36	31.3	23	US-07-841-591A-14	Sequence 14, Appl
5	36	31.3	23	PCT-US93-02034-14	Sequence 14, Appl
6	34	29.6	21	US-08-726-306A-15	Sequence 15, Appl
7	33	28.7	14	US-09-217-228-16	Sequence 16, Appl
8	33	28.7	17	US-09-082-358B-30	Sequence 30, Appl
9	32	27.8	19	PCT-US91-02942-67	Sequence 67, Appl
10	32	27.8	20	US-09-579-664B-27	Sequence 27, Appl
11	32	27.8	21	US-08-557-309B-58	Sequence 58, Appl
12	32	27.8	21	US-08-929-414-11	Sequence 11, Appl
13	32	27.8	21	US-08-834-306-56	Sequence 56, Appl
14	32	27.8	21	US-08-834-306-58	Sequence 58, Appl
15	32	27.8	21	US-08-993-674A-58	Sequence 58, Appl
16	32	27.8	21	US-08-993-674A-58	Sequence 58, Appl
17	32	27.8	21	US-09-256-976-56	Sequence 56, Appl
18	32	27.8	21	US-09-256-976-58	Sequence 58, Appl
19	32	27.8	22	US-07-916-034-22	Sequence 22, Appl
20	31.5	27.4	21	US-08-557-309B-60	Sequence 60, Appl
21	31.5	27.4	21	US-08-929-414-12	Sequence 12, Appl
22	31.5	27.4	21	US-08-834-306-59	Sequence 59, Appl
23	31.5	27.4	21	US-08-993-674A-59	Sequence 59, Appl
24	31.5	27.4	21	US-09-256-976-59	Sequence 59, Appl
25	31	27.0	10	5240848-2	Patent No. 5240848
26	31	27.0	12	US-08-742-243-4	Sequence 4, Appli
27	31	27.0	14	5219739-39	Patent No. 5219739

28	31	27.0	16	2	US-08-660-592-11	Sequence 11, Appl
29	31	27.0	16	2	US-08-437-013-16	Sequence 16, Appl
30	31	27.0	16	4	US-09-275-508A-16	Sequence 16, Appl
31	31	27.0	17	4	US-09-244-583-10	Sequence 10, Appl
32	31	27.0	19	2	US-08-660-592-10	Sequence 10, Appl
33	30.5	26.5	20	3	US-09-247-527-18	Sequence 18, Appl
34	30	26.1	12	1	US-08-433-854-25	Sequence 25, Appl
35	30	26.1	12	1	US-08-174-745A-25	Sequence 25, Appl
36	30	26.1	12	2	US-08-195-947-25	Sequence 25, Appl
37	30	26.1	12	2	US-08-433-885-25	Sequence 25, Appl
38	30	26.1	12	2	US-08-433-908B-25	Sequence 25, Appl
39	30	26.1	12	3	US-08-410-614-25	Sequence 25, Appl
40	30	26.1	16	1	US-08-333-565-26	Sequence 26, Appl
41	30	26.1	16	1	US-07-942-245-105	Sequence 105, App
42	30	26.1	16	2	US-08-661-479-26	Sequence 26, Appl
43	30	26.1	18	3	US-08-105-454-6	Sequence 6, Appli
44	30	26.1	18	4	US-09-822-624-17	Sequence 17, Appl
45	30	26.1	19	4	US-09-050-739-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1  
US-08-053-451B-171  
; Sequence 171, Application US/08053451B  
; Patent No. 5955584  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Francis W.  
; APPLICANT: Ditlow, Charles C.  
; APPLICANT: Calenoff, Emanuel  
; TITLE OF INVENTION: ATHEROSCLEROTIC PLAQUE SPECIFIC  
; TITLE OF INVENTION: ANTIGENS, ANTIBODIES THERETO, AND USES THEREOF  
; NUMBER OF SEQUENCES: 176  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/053,451B  
; FILING DATE: 26-APR-1993  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Halluin, Albert P.  
; REGISTRATION NUMBER: 25,227  
; REFERENCE/DOCKET NUMBER: 7606-033-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-854-3660  
; TELEFAX: 415-854-3694  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 171:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-053-451B-171

Query Match 33.0%; Score 38; DB 2; Length 20;  
Best Local Similarity 50.0%; Pred. No. 31;  
Matches 10; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

QY 1 GTAPAAEGAGA--EVKRAA 18  
| : | ||| | : ||| |  
| : | ||| | : ||| |

```

Db      1  GESPLTFGAGTKLEKRAA 20

RESULT 2
US-09-319-570A-4
; Sequence 4, Application US/09319570A
; Patent No. 6537760
; GENERAL INFORMATION:
; APPLICANT: BERGMANN, ANDREAS
; APPLICANT: STRUCK, JOACHIM
; TITLE OF INVENTION: RECEPTOR BINDING ASSAY FOR DETECTING TSH RECEPTOR
; TITLE OF INVENTION: AUTOANTIBODIES AND REAGENT KIT FOR CARRYING OUT SUCH A
; TITLE OF INVENTION: RECEPTOR BINDING ASSAY
; FILE REFERENCE: 011377-0260244
; CURRENT APPLICATION NUMBER: US/09/319,570A
; CURRENT FILING DATE: 1999-08-04
; PRIOR APPLICATION NUMBER: DE 196 51 093.7
; PRIOR FILING DATE: 1996-12-09
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-319-570A-4

Query Match      31.3%; Score 36; DB 4; Length 21;
Best Local Similarity 47.1%; Pred. No. 64;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY      8  GAGAEVVKRASAEAKQAF 24
      ||| ||| ||| |||
Db      5  GFGQELKNPQETLQAF 21

RESULT 3
US-09-236-423-1
; Sequence 1, Application US/09236423
; Patent No. 6395875
; GENERAL INFORMATION:
; APPLICANT: Freimuth, Paul
; TITLE OF INVENTION: RECOMBINANT SOLUBLE ADENOVIRUS RECEPTOR
; FILE REFERENCE: BNL-2007
; CURRENT APPLICATION NUMBER: US/09/236,423
; CURRENT FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: CLONING VECTOR
; OTHER INFORMATION: ENCODED SEQUENCES
US-09-236-423-1

Query Match      31.3%; Score 36; DB 4; Length 22;
Best Local Similarity 47.4%; Pred. No. 68;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      4  PAAGAGAEVVKRASAEAKQ 22
      ||| ||| ||| |||
Db      4  PAANKARKEAELAAATAEQ 22

RESULT 4
US-07-841-591A-14
; Sequence 14, Application US/07841591A
; Patent No. 5900476
; GENERAL INFORMATION:
; APPLICANT: Ruggeri, Zaverio M. and
; APPLICANT: Ware, Jerry, inventors
; APPLICANT: on behalf of The Scripps Research
; APPLICATION NUMBER: PCT/US93/02034
; FILING DATE: 19930223

```

```

; APPLICANT: Institute
; TITLE OF INVENTION: Therapeutic Domains of
; TITLE OF INVENTION: von Willebrand Factor
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute
; STREET: 10666 No. 5900476th Torrey Pines Road
; CITY: La Jolla
; STATE: California
; COUNTRY: United States
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.2 megabyte 5 1/4" floppy
; COMPUTER: AST Bravo 386SX
; OPERATING SYSTEM: MS DOS version 3.2
; SOFTWARE: WordPerfect 5.1 conv. to ASCII
; CURRENT APPLICATION DATA: US/07/841,591A
; APPLICATION NUMBER: US/07/841,591A
; FILING DATE: 26-Feb-92
; CLASSIFICATION:
; PRIOR APPLICATION DATA: This is a c-i-p of
; APPLICATION NUMBER: PCT/US91/07756
; FILING DATE: 17-Oct-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Barron, Alexis
; REGISTRATION NUMBER: 22,702
; REFERENCE/DOCKET NUMBER: P16,633-H
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 923-4466
; TELEFAX: (215) 923-2189
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23
; TYPE: Amino acid
; STRANDEDNESS:
; TOPOLOGY: Linear
US-07-841-591A-14

Query Match      31.3%; Score 36; DB 2; Length 23;
Best Local Similarity 47.4%; Pred. No. 71;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY      4  PAAGAGAEVVKRASAEAKQ 22
      ||| ||| ||| |||
Db      5  PAANKARKEAELAAATAEQ 23

RESULT 5
PCT-US93-02034-14
; Sequence 14, Application PC/TUS9302034
; GENERAL INFORMATION:
; APPLICANT: Ruggeri, Zaverio M. and
; APPLICANT: Ware, Jerry, inventors
; APPLICANT: on behalf of The Scripps Research
; APPLICANT: Institute
; TITLE OF INVENTION: Therapeutic Domains of
; TITLE OF INVENTION: von Willebrand Factor
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute
; STREET: 10666 North Torrey Pines Road
; CITY: La Jolla
; STATE: California
; COUNTRY: United States
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 2 megabyte 3.25" floppy
; COMPUTER: AST Bravo 386SX
; OPERATING SYSTEM: MS DOS version 3.2
; SOFTWARE: WordPerfect 5.1 conv. to ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/02034
; FILING DATE: 19930223

```

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA: This is a c-i-p of
; APPLICATION NUMBER: PCT/US91/07756
; FILING DATE: 17-Oct-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Barron, Alexis
; REGISTRATION NUMBER: 22,702
; REFERENCE/DOCKET NUMBER: P16,633-H PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 923-4466
; TELEFAX: (215) 923-2189
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: Linear
; PCT-US93-02034-14

Query Match 31.3%; Score 36; DB 5; Length 23;
Best Local Similarity 47.4%; Pred. No. 71;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 PAAGAGAEVVKRAAEAKQ 22
DB 5 PAANKARKEAEAAATAEQ 23

RESULT 6
US-08-726-306A-15
; Sequence 15, Application US/08726306A
; Patent No. 5958684
; GENERAL INFORMATION:
; APPLICANT: van Leeuwen, Frederik Willem
; APPLICANT: Burbach, Johannes Peter Henri
; APPLICANT: Grosveld, Franklin G.
; TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
; NUMBER OF SEQUENCES: 189
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Witcoff, Ltd.
; STREET: 1 Financial Center
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726,306A
; FILING DATE: 02-Oct-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 95/20080.4
; FILING DATE: 02-Oct-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,822
; FILING DATE: 01-Jan-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Ph.D., Kathleen M.
; REGISTRATION NUMBER: 34,380
; REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 345-9100
; TELEFAX: (617) 345-9111
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide

```

```

US-08-726-306A-15
Query Match 29.6%; Score 34; DB 2; Length 21;
Best Local Similarity 61.5%; Pred. No. 1.3e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 1 GTAPAA--EGAGA 11
DB 5 GTAPAGRLDGSQA 17

RESULT 7
US-09-217-228-16
; Sequence 16, Application US/09217228
; Patent No. 6323178
; GENERAL INFORMATION:
; APPLICANT: Butler, Jon P.
; APPLICANT: Hale, John E.
; APPLICANT: Heath Jr., William F.
; APPLICANT: Schoner, Brigitte E.
; APPLICANT: Heiman, Mark L.
; APPLICANT: Becker, Gerald W.
; APPLICANT: Varshavsky, Alexander D.
; TITLE OF INVENTION: Beta-lipotropin and Uses Thereof
; FILE REFERENCE: X-12139
; CURRENT APPLICATION NUMBER: US/09/217,228
; CURRENT FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-217-228-16

Query Match 28.7%; Score 33; DB 4; Length 14;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 PAAGAGAE 12
DB 2 PADDGAGAQ 10

RESULT 8
US-09-082-358B-30
; Sequence 30, Application US/09082358B
; Patent No. 6469153
; GENERAL INFORMATION:
; APPLICANT: Goff, Stephen P.
; APPLICANT: Li, Xingqiang
; TITLE OF INVENTION: EIP-1, EIP-3 GENES, ENVELOPE-INTERACTING PROTEINS,
; FILE REFERENCE: EIP-1, and EIP-3
; FILE REFERENCE: 0575/54804
; CURRENT APPLICATION NUMBER: US/09/082,358B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 17
; TYPE: PRT
; ORGANISM: murine
; US-09-082-358B-30

Query Match 28.7%; Score 33; DB 4; Length 17;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 EGAGAEVVKRA 16
DB 1 DGAGARVLRA 10

```

## RESULT 9

PCT-US91-02942-67  
; Sequence 67, Application PC/TUS9102942  
; GENERAL INFORMATION:  
; APPLICANT: ROTHLEIN, ROBERT  
; APPLICANT: ADAIR, JOHN R  
; APPLICANT: ATHWAL, DILJEET S  
; TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY  
; NUMBER OF SEQUENCES: 102  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
; STREET: 1225 Connecticut Ave. NW Suite 300  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/02942  
; FILING DATE: 19910429  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER:  
; FILING DATE: 27-APR-1990  
; APPLICATION NUMBER: GB 9009549.8  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FOX, SAM L  
; REGISTRATION NUMBER: 30,353  
; REFERENCE/DOCKET NUMBER: 1011.0586600  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 466-0800  
; TELEFAX: (202) 833-8716  
; INFORMATION FOR SEQ ID NO: 67:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 19 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US91-02942-67

Query Match 27.8%; Score 32; DB 5; Length 19;  
Best Local Similarity 46.2%; Pred. No. 2.2e+02;  
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 9 AGAEVKRASAEAK 21  
:|||||: :  
Db 2 SGAEVKPGSSVK 14

## RESULT 10

US-03-579-664B-27  
; Sequence 27, Application US/09579664B  
; Patent No. 6514719  
; GENERAL INFORMATION:  
; APPLICANT: Immunex Corporation  
; APPLICANT: Bird, Timothy A.  
; APPLICANT: Virca, G. Duke  
; APPLICANT: Martin, Unja  
; APPLICANT: Anderson, Dirk M.  
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES  
; FILE REFERENCE: 2923-A  
; CURRENT APPLICATION NUMBER: US/09/579,664B  
; CURRENT FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 27  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:

; OTHER INFORMATION: peptide  
US-09-579-664B-27

Query Match 27.8%; Score 32; DB 4; Length 20;  
Best Local Similarity 47.4%; Pred. No. 2.3e+02;  
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 5 AAEAGAEVKRASAEAKQA 23  
:|||||: :  
Db 2 AAAREPAEPSPFAAEAGA 20

## RESULT 11

US-08-557-309B-58  
; Sequence 58, Application US/08557309B  
; Patent No. 5916572  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Houghton, Raymond L.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T  
; NUMBER OF SEQUENCES: 69  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SEED and BERRY LLP  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/557,309B  
; FILING DATE: 14-NOV-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 210121.422  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 622-4900  
; TELEFAX: (206) 682-6031  
; INFORMATION FOR SEQ ID NO: 58:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-557-309B-58

Query Match 27.8%; Score 32; DB 2; Length 21;  
Best Local Similarity 47.1%; Pred. No. 2.4e+02;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 7 EGAGAEVKRASAEAKQA 23  
:|||||: :  
Db 1 KAAAFAPAKAAAPAKAA 17

## RESULT 12

US-08-929-414-11  
; Sequence 11, Application US/08929414  
; Patent No. 5942403  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Houghton, Raymond  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION  
; FEATURE: OF T. CRUZI INFECTION

NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/929,414  
FILING DATE: 15-SEP-1997  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.406C1  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-929-414-11

Query Match 27.8%; Score 32; DB 2; Length 21;  
Best Local Similarity 47.1%; Pred. No. 2.4e+02;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 7 EGAGAEVVKRASAEAKQA 23  
Db 1 KAAAPAKAAAPAKAA 17

RESULT 13  
US-08-834-306-56  
Sequence 56, Application US/08834306  
Patent No. 6054135  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Houghton, Raymond L.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/834,306  
FILING DATE: 15-APR-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.422C1  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-834-306-56

Query Match 27.8%; Score 32; DB 3; Length 21;  
Best Local Similarity 47.1%; Pred. No. 2.4e+02;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 7 EGAGAEVVKRASAEAKQA 23  
Db 1 KAAAPAKAAAPAKAA 17

RESULT 14  
US-08-834-306-58  
Sequence 58, Application US/08834306  
Patent No. 6054135  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Houghton, Raymond L.  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED and BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/834,306  
FILING DATE: 15-APR-1997  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.422C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-08-834-306-58

Query Match 27.8%; Score 32; DB 3; Length 21;  
Best Local Similarity 47.1%; Pred. No. 2.4e+02;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 7 EGAGAEVVKRASAEAKQA 23  
Db 1 KAAAPAKAAAPAKAA 17

RESULT 15  
US-08-993-674A-56  
Sequence 56, Application US/08993674A

```

; Patent No. 6228372
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Smith, John M.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND PREVENTION OF T
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/993,674A
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
US-08-993-674A-56

```

```

Query Match          27.8%; Score 32; DB 3; Length 21;
Best Local Similarity 47.1%; Pred. NO. 2.4e+02;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
Qy      7 EGAGAEVVKASAEAKQA 23
Db      1 KAAAPAKAAAPAKAA 17

```

Search completed: October 7, 2004, 17:58:14  
Job time : 23 secs